

مطبوعات
مكتبة الملك فهد الوطنية
السلسلة الأولى
(٥٤)

أساسيات ما وراء البيانات

لاختصاصيي المكتبات والمعلومات

تأليف
بريسيلا كابلن

ترجمة وتعليق
د. هاشم فرحات

الرياض
١٤٢٨هـ - ٢٠٠٧م

الإهـــــــداء

إلى ثلاثة قناديل ... لكل منهم في النفس محبة ووفاء وتقدير

إلى والديَّ الكريمين ...

تغمدهما المولى ـ سبحانه وتعالى ـ بواسع رحمته وأسكنهما فسيح جناته

وإلى :

العالم الجليل صاحب الفكر الرائد في مجال هذا الكتاب وموضوعه ...
الأستاذ الدكتور سعد الهجرسي

وإلى :

الباحثين في مجال المكتبات والمعلومات ... الذين تتعلق في أعناقهم أمانة النهوض بهذا التخصص الحيوي ، ومسؤولية تثبيت أركانه ... في ظل مجتمعات تعالت فيها صيحات غير مسؤولة ـ عن قصد أو عن جهل ـ لتهميش دوره وتشويه هويته ....

المترجم

المحتويات

قائمة الأشكال التوضيحية

الموضوع **الصفحة**

مقدمة المترجم

عن القضية وأبعادها

الحمد لله رب العالمين ، والصلاة والسلام على النبي الأمين ، وعلى آله وصحبه أجمعين، وبعد:

فمنذ بدايات عقد الستينات من القرن العشرين الميلادي المنصرم تحيط بمجال المكتبات بشكل عام، ومجال تنظيم المعلومات بشكل خاص، تطورات تقنية كثيرة ومتنوعة، وقد حاول المترجم في دراسة سابقة رصد أهم تلك التطورات (١)، وخاصة ما ارتبط منها بمجال تنظيم المعلومات، واستكشاف انعكاسات تلك التطورات وتأثيراتها في المجال، وقد تبين أن هذه التطورات كانت لها بالفعل آثار قوية دفعت لردة فعل إيجابية أقوى من قبل القوى الفاعلة في هذا المجال، وأقصد بالقوى الفاعلة هنا المؤسسات الأكاديمية والجمعيات العلمية؛ حيث أعدت الأولى عدتها لمراجعة مقرراتها ومناهجها الدراسية والتأهيلية لتتواكب مع هذه التطورات، وتبنت الثانية عقد اللقاءات والمؤتمرات لطرح انعكاسات التطورات الحديثة على جبهة البحث والمناقشة التماساً لسبل تتواءم مع طبيعة هذه التطورات ، وتفيد من إيجابياتها، ثم تستثمر إمكاناتها .

كما تبين كذلك من خلال مراجعة تلك التطورات أن من أكثر القضايا

(١) هاشم فرحات، ومبارك سعد سليمان. التطورات الحديثة في مجال تنظيم المعلومات وانعكاساتها على تعليم الفهرسة في أقسام علوم المكتبات والمعلومات العربية. مجلة المكتبات والمعلومات العربية ، س ٢٦، ع١ (يناير ٢٠٠٦م)، ص٥٩-٩٤ .

التي ألقت بظلالها على مجريات الأمور في عالم المكتبات تلك التي تتصل بقضايا تنظيم مصادر المعلومات الإلكترونية المتاحة عن بعد بشكل خاص، تلك الفئة الجديدة التي فرضتها الإنترنت على الساحة مع بدايات عقد التسعينات من القرن العشرين الميلادي عندما اتسع مجالها وتداعت جهود استثمار إمكاناتها في نشر الإنتاج الفكري وإتاحته عبر جنباتها مترامية الأطراف، وعلى أثيرها غير المحدود. وقد تبين كذلك أن المؤسسات الأكاديمية تفاعلت مع هذه القضية تفاعلاً إيجابياً انعكس على برامج تأهيلها للمكتبيين وتدريبهم على التعامل مع هذا الوافد الجديد، وإدارته إدارة منهجية تعمل على تدبير أساليب الحصول عليه وتنظيمه وإعداده فنياً بما يكفل إتاحته للمستفيدين، ومن ثم تيسير سبل الإفادة منه.

ومع ارتفاع معدلات إنتاج هذا النمط من مصادر المعلومات، وارتفاع معدلات نشره وإتاحته، وتنامي أبعاد استخدامه وتداوله والإفادة منه، بدت الحاجة، وبشكل أكثر حدة من ذي قبل، إلى أدوات تساعد على ضبطه وتنظيمه وإدارته، من هنا كان على اختصاصيي المكتبات والمعلومات أن يبدوا مزيداً من الاهتمام ، وأن تتضافر جهودهم، بل أن تتضاعف، في هذا الاتجاه، ولكن مما لفت الانتباه أن جهود غيرهم - أي من غير المكتبيين - وخاصة أولئك المعنيين بتلك الفئة من المصادر إنتاجاً أو نشراً أو إفادة ، كانت أكثر تضاعفاً ، بصرف النظر عن مدى تقنينها، وخصوصاً أن هناك مؤشرات تفيد بأن من أقوى الدوافع التي أدت إلى تضاعف تلك الجهود وتفاقمها هو زعزعة الثقة بجهود المكتبيين، بل التوجس خيفة من عدم قدرة تلك الجهود على تحقيق ما يصبو إليه كل المعنيين بتلك المصادر، وخاصة

الحريصين على تداولها والإفادة منها.

عن المصطلح وأبعاده :

عند هذه النقطة، ومع وجود مصطلحات كثيرة راسخة ومتداولة في عالم المكتبات تتصل بإجراءات تنظيم مصادر المعلومات وإعدادها فنياً، مثل مصطلح "الفهرسة" في سياقها التقليدي، وما تعرف به الأداة الناتجة عن هذه العملية - أعني بطاقة الفهرسة، ومصطلح الفهرسة المقروءة آلياً، وما تعرف به الأداة الناتجة عن هذه العملية كذلك ، أعني الفهرسة المقروءة آلياً ، أو تسجيلة مارك وغيرها، مع كل ذلك بدا في الأفق مصطلح جديد يعبر عن رؤية غير المكتبيين لعملية الفهرسة من ناحية، ويعكس توجهات جهودهم تجاه تنظيم هذا النمط الجديد من المصادر من ناحية أخرى، ذلكم هو مصطلح: "metadata" - الذي يعد بمنزلة إطار عام تنضوي تحت مظلته جميع جهود تنظيم ووصف وفهرسة مصادر المعلومات الإلكترونية بشكل عام ، سواء المتاح منها على وسائط إلكترونية حديثة، أو المصادر الرقمية المتاحة عن بعد من خلال الإنترنت.

ولقد لفت انتباهي كمتابع لهذا الموضوع أن اختصاصيي المكتبات، ومع تحفظات كثير منهم على ما وراء هذا المصطلح، وقناعتهم بأن جوهر ما يطمح إلى تحقيقه غير المكتبيين أقرب ما يكون من الحرث لأرض محروثة، وليس للجديد حظ منه؛ ذلك انطلاقاً من قناعتهم بأنهم هم السباقون إلى هذه الأرض التي زرعوا فيها جهودهم وترسخت فيها أقدامهم منذ مئات السنين، واستدلوا على ذلك بما يتداول من قواعد الفهرسة وتقنياتها وأدواتها متتالية الطبعات

والإصدارات .

في مقابل هذه القناعات ومع وجاهتها لم يكن لهؤلاء المكتبيين من حيلة لرفض ذلكم المصطلح (أعني metadata) ، وإنما تفاعلوا معه، وباتوا يتدارسون مبادرات وجهود أندادهم ممن طرحوه على ساحتهم ، بل وجدوا في كثير من تلك الجهود ما يدعوهم للتأمل ومراجعة الثوابت وتقييمها في ضوء هذا الواقع الذي فرض عليهم ، بل أكثر من ذلك فقد دعاهم ذلك إلى إعادة النظر في ممارساتهم وجهودهم تجاه الوافد الجديد - أعني المصادر الإلكترونية - وقد اتخذت نتائج هذه التأملات والمراجعات بعداً إيجابياً يسعى للمحافظة على الثوابت من جهة، ويتفاعل مع الجديد من جهة أخرى، وقد تجسدت تلك التوجهات الإيجابية في صور مختلفة، كان من أبرزها على الساحة التعليمية إعادة النظر في جميع مقررات الفهرسة التي تدرّس في أقسام علوم المكتبات والمعلومات ومحاولة تطويرها، وقد اتخذت اتجاهات التطوير ثلاثة أبعاد أساسية : تمثل أولها في إعادة النظر في المقررات التقليدية كالفهرسة الوصفية والتصنيف والفهرسة الموضوعية، واستبدالها بمقررات حديثة تتم في سياقها معالجة كثير من القضايا الحديثة المرتبطة بهذا الموضوع ، وغالباً ما ترد هذه المقررات الحديثة تحت اسم " تنظيم المعلومات"، وتمثل ثانيها في تضمين تلك البرامج مقررات جديدة مستقلة تماماً عن فهرسة مصادر أو موارد الإنترنت وما وراء البيانات، أما ثالثها فتمثل في تخصيص مقررات مستقلة لمعالجة قضايا ما وراء البيانات بدءاً من المقدمات التي تهدف إلى استيعاب أساسياتها وعلاقاتها بالفهرسة وغيرها من العمليات الفنية، وانتهاءً بالفنيات التي تهدف إلى دراسة وبحث قضايا

دقيقة كقضايا التشغيل البيني ونقاط المقابلة أو التحويل البيني Crosswalking، مروراً بالمعالجات التفصيلية لخطط ما وراء البيانات ومعاييرها.

أما على الساحة البحثية فقد تجسدت جهودهم في محاور متعددة ، تمثل بعضها في عقد اللقاءات العلمية المشتركة مع المبادرين باختراع ذلك المصطلح، وتمثل بعضها الآخر في مشاركة هؤلاء في بعض مبادراتهم وجهودهم ـ ولعل أبرز ما يشار إليه في هذا الصدد ما تم في من تضافر للجهود بين الفريقين في طرح مبادرة دبلن كور لما وراء البيانات، بل هناك توجه ثالث تمثل في طرح مبادرات جديدة في هذا الصدد كتلك المبادرات التي تبنتها مكتبة الكونجرس لتطوير معايير خاصة بما وراء البيانات.

عن الكتاب ودوافع اختياره :

إذا كانت جهود ومبادرات خطط ومعايير ما وراء البيانات بمعناها الضيق الذي يتمحور حول المصادر الإلكترونية قد مضى عليها ما يزيد على عقد من الزمان، على اعتبار أن أولى تلك المبادرات - مبادرة دبلن كور - أطلقت مع منتصف التسعينات من القرن العشرين الميلادي ، وإذا كانت التطورات التي شهدتها مناهج ومقررات البرامج التأهيلية في علم المكتبات والمعلومات في هذا الصدد تحديداً قد تزامنت تقريباً مع تلك المبادرات واستمرت حتى هذه الآونة، فقد لفت انتباهي مرة أخرى تعدد الإنتاج الفكري - الأجنبي بالطبع وليس العربي - حول قضايا "ما وراء البيانات" وارتفاع معدلات نشر هذا الإنتاج ، وتعدد أشكاله وفئاته، وقد قدره بعض الباحثين في سياق مراجعة علمية حديثة بخمس مئة وسبعين عملاً

نـشرت خـلال الفتـرة مـن ١٩٩٧ ـ ٢٠٠٤ م[١]، ولكـن مـن أهـم مـا يمكـن تسجيله على هذا الإنتاج هو غلبة مقالات الدوريات ؛ إذ تصل نسبتها إلى ما يقرب من ثلثي هذا الإنتاج، في مقابل ذلك ثمة انخفاض في نسبة الكتب المنشورة؛ إذ لم تتعد نسبتها ١%، وفي ذلك إشارة واضحة إلى خلو الساحة العلمية من عمل شامل يلملم أطراف هذا الموضوع ويجمع شتاته، بل من أكثر الأمور التي لفتت انتباه المترجم أن عدد الكتب التي نشرت كاملة عن الموضوع حتى كتابة هذه المقدمة لم يتجاوز أصابع اليدين.

وإذا كان هذا هو واقع الحال بالنسبة للإنتاج الفكري الأجنبي، فلم يكن واقع الساحة العربية بأحسن حالاً، إن لم يكن بالفعل ذا سمات ليس لها من وجه المقارنة أي نصيب؛ حيث أظهرت المراجعة الأخيرة التي أشير إليها في الفقرة السابقة أن مجمل ما نشر من إنتاج عربي في هذا الموضوع لم يتعد ستة وعشرين عملاً، ليس من بينها كتاب واحد ، باستثناء أحد الكتب المنشورة حديثًا[٢]، وهو كتاب على أي حال لم يخصص لمعالجة هذه القضية معالجة مستقلة، وإنما جاءت معالجته لها منجمة بين ثناياه، وفي مواضع متفرقة.

(١) زين عبد الهادي. فهرسة مصادر الإنترنت: مراجعة علمية للإنتاج الفكري. ندوة فهرسة مصادر الإنترنت واستخدام معايير الميتاديتا ودبلن كور .- القاهرة: المنظمة العربية للتنمية الإدارية، ٤-٢٠٠٤/٧/٨م. محمود عبد الكريم الجندي. ما وراء البيانات: مراجعة علمية للإنتاج الفكري، ٢٠٠٥ (قيد النشر).

(١) آرمز، وليم. المكتبات الرقمية ؛ ترجمة هاشم فرحات، وجبريل عريشي .- الرياض: مكتبة الملك فهد الوطنية، ١٤٢٧هـ /٢٠٠٦م.

أمام هذه الشواهد، وفي ظل حاجتنا الفعلية - في العالم العربي - إلى عمل شامل يجلي لنا ما خفي عن حقيقة هذا القضية، التي أحسبها محورية لا فرعية، ويظهر لنا ما نجهل عنها ويعالجها في سياق واضح ومحدد من جهة، ويدعم رؤية كان المترجم قد طرحها في سياق عمل سابق من جهة أخرى ، تلك الرؤية التي كان مفادها إعادة النظر في مقررات الفهرسة في أقسام علوم المكتبات والمعلومات العربية وتبني تدريس مقررات مستقلة أو متكاملة تعالج قضايا ما وراء البيانات(١)، بدأت رحلة التفكير في عمل يسعى إلى تلبية تلك الحاجة، لكن سرعان ما وجدتني متوقفاً عند التفكير في طبيعة هذا العمل، ومتردداً في الشكل الذي ينبغي أن يظهر به، فهل من الأفضل أن يكون مترجماً أم مؤلفاً؟ ، وبعد مدة غير قصيرة من التفكير رجحت كفة الترجمة على مشقتها وتبعاتها، على أمل أن ترجأ الرغبة الثانية - أي التأليف - إلى حين استكمال كل ما يطمح المرء إلى أن يجمعه حول هذا الموضوع الجديد.

وبعد الاستقرار على خيار الترجمة جاءت رحلة المفاضلة التي لم تدم طويلاً في ظل محدودية البدائل وصعوبة الحصول عليها جميعاً؛ حيث تبين من خلال مراجعة كثير من الدراسات الحديثة حول الموضوع أن هذا العمل الذي بين أيدينا يعد من أكثر الأعمال التي يستشهد بها في هذا الصدد، بل

(١) جاء ذلك تحديداً بعنوان "نحو تصور مقترح لتعليم الفهرسة في أقسام علوم المكتبات في المنطقة العربية"، ص ٨١-٩٠ ، في: هاشم فرحات، ومبارك سعد سليمان. المصدر السابق.

أكاد أجزم بأنه لم يكد يخلو عمل كتب عن موضوع ما وراء البيانات من قريب أو بعيد إلا وقد اعتمد عليه وتواترت مرات الاستشهاد به أو الإشارة إليه، ذلك لسببين أساسيين: يتمثل أولهما في أن هذا الكتاب يعد أول كتاب يتحدث عن الموضوع بشمولية، سواء من حيث الأساسيات أم من حيث تناوله جميع خطط ومعايير ما وراء البيانات العامة أو شائعة الاستخدام، حيث جاء الكتاب في قسمين: يتناول أولهما أساسيات ما وراء البيانات وجاء في خمسة فصول، ويتناول ثانيهما خطط ما وراء البيانات، وجاء في ثلاثة عشر فصلاً، خصصت العشرة الأولى منها للتعريف بمعايير ما وراء البيانات المتداولة الآن في الميدان، بحيث يتناول كل فصل منها معياراً مستقلاً ومتخصصاً في مجال محدد، ثم يختتم هذا القسم بالحديث عن الأنواع الأخرى لما وراء البيانات غير ما وراء البيانات الوصفية التي عولجت بشكل وافٍ في القسم الأول، وقد خصصت لذلك الفصول الثلاثة الأخيرة، حيث يتناول أحدها ما وراء البيانات الإدارية، ويتناول الثاني ما وراء البيانات البنائية ، ويتناول الثالث ما وراء البيانات الخاصة بالحقوق، وبذلك تكتمل منظومة هذا الكتاب التي جاءت في ثمانية عشر فصلاً، يبلغ عدد صفحاتها مئة وثنتين وتسعين صفحة من القطع الكبير، تخللها سبعة وعشرون شكلاً توضيحياً إما لنصوص مقتبسة من بعض المعايير أو لصفحات من الإنترنت لبعض التطبيقات، فضلاً عن عشرات النماذج لصيغ تسجيلات ما وراء البيانات في صيغها المختلفة التي لم تكد تخلو منها صفحة من صفحات الكتاب.

أما السبب الثاني الذي دفعني لاختيار هذا الكتاب، فيتمثل في خبرة

المؤلفة ومؤهلاتها؛ فبريسيلا كابلن، مؤلفة هذا الكتاب، لديها من المؤهلات والخبرات المهنية ما يرشحها لأن تكون من الثقات في هذا المجال، فهي - وكما عرفت بنفسها في خاتمة الكتاب، وكما يتسطر في صفحتها على الإنترنت[1] - ممن لهم باع طويل في مجال تقنيات المعلومات وارتباطها بمجال المكتبات والمعلومات بشكل عام ، وفي مجال المكتبات الرقمية ومشروعات الرقمنة بشكل خاص، وفي مجال ما وراء البيانات بشكل أخص . أما من حيث خلفيتها العلمية وتأهيلها الأكاديمي والوظائف التي عملت فيها، فمؤلفة الكتاب حاصلة على الماجستير في علم المكتبات M. L. S . من جامعة نورث كارولينا North Carolina في تشابل هل Chapel Hill. وهي تعمل الآن مديرة مساعدة لخدمات المكتبة الرقمية في مركز فلوريدا لأتمتة المكتبات في جينسفيل Gainesville بولاية فلوريدا، وكانت قد عملت في الوظيفة نفسها - أي مديرة مساعدة لنظم المكتبات في مكتبة جامعة شيكاغو، وكذلك رئيسة لقسم تطوير النظم في مكتب نظم المعلومات في مكتبة جامعة هارفارد، ومن خبراتها كذلك أنها ترأست إحدى اللجان الاستشارية لمارك التابعة للجمعية الأمريكية للمكتبات والمعروفة بلجنة ماربي MARBI التي تعنى بدراسة واقتراح عناصر بيانات وحقول مارك الخاصة بالمصادر الإلكترونية، والمجموعة الاستشارية لمارك، كما رأست لجنة تطوير المعايير في المنظمة الوطنية لمعايير المعلومات (NISO)، فضلاً عن أنها عضو في اللجنة الاستشارية لدبلن كور. وأخيراً تعد المؤلفة

---

(١) يمكن مراجعة صفحة المؤلفة على الإنترنت على الموقع التالي:
http: //www. fcla. edu/~pcaplan/#pubs

من المكثرين في المجال وتحديداً فيما يتصل بالكتابة عن قضايا المكتبات الرقمية وما وراء البيانات، حيث نشرت ما يقرب من عشرين بحثاً في هذا المجال.

عن المصطلح في سياقه العربي :

وأخيراً تبقى الإشارة إلى قضية شائكة، وقد كثر الحديث عنها في سياقات مختلفة، ولا أريد الدخول في تفاصيلها في هذا السياق، ولكن حسبي الإشارة إلى بعض ملابساتها وخاصة ما يتصل منها بموضوع هذا الكتاب وعنوانه، فالقضية أولاً هي قضية ترجمة المصطلحات الجديدة، ومدى توحيد الجهود في هذا الصدد، أما ما يختص منها بعنوان هذا الكتاب الذي بين أيدينا فهو في حقيقة الأمر يجسد جوهر القضية في لحمتها وسداها، فمحور الكتاب يدور حول موضوع "metadata" - ويبدو واضحاً من البداية أنه مصطلح حديث، وله دلالته في اللغة الإنجليزية، وقد تتبعت كثيراً من تعريفاته، فوجدت جلها يكاد يجمع على أن المقصود بكلمة "metadata" – التي تكتب في شكلين آخرين: أولهما: "meta data"، وثانيهما "meta-data" - هو "البيانات عن البيانات data about data"، ذلك أن البادئة "meta" في سياق مصطلحات علوم الحاسب تقابل كلمة " about " بالإنجليزية ، وهي ما تقابل بالعربية في أبسط معانيها كلمة "عن" ، ومن ثم فهي كلمة مركبة لها دلالتها الواضحة، ولكن عند ترجمتها إلى العربية يكون من غير المستساغ القول "بيانات عن البيانات .

من هنا تعددت جهود ترجمتها، واجتهد غير واحد من الباحثين وخاصة

من تناول هذا الموضوع أو اهتم به ، بل ألقى كل واحد منهم بدلوه في هذا الصدد حتى وصلت جهود هؤلاء - كما أشارت المراجعة التي سبقت الإشارة إليها(١) - إلى وضع نحو ثلاث عشرة ترجمة عربية لهذا المصطلح، هي: "ما وراء البيانات"، و"ما بعد البيانات"، و"بيانات البيانات"، و"بيانات الوصف المدمجة في صفحات الإنترنت"، و"وصائف البيانات"، و"واصفات البيانات"، و"البيانات الفارقة"، و"ما فوق البيانات"، و"البيانات الواصفة"، و"البيانات الخلفية"، و"بيانات عن بيانات"، و"ميتاديتا"، و"ميتاداتا" ، وفضلاً عن هذه البدائل الثلاثة عشر، عنّ لي في وقت من الأوقات أن أضيف ثلاثة بدائل أخرى رأيت قرب دلالتها من الموضوع، وهي: "البيانات الخفية أو المخفية"، و"بيانات الوصف الخفية"، و"البيانات المتعدية".

ومع تقديري لجميع هذه الجهود وقناعتي بأن أياً منها يعكس رؤية كل باحث وفهمه طبيعة مثل هذه البيانات، وأن أياً منها قابل للنقد والمراجعة غير أن ذلك يخرج عن هذا السياق ، فقد عدلت عنها جميعاً في الوقت الحاضر، برغم قناعتي بقدرتها على التعبير عن طبيعة هذه البيانات وهدفها؛ فهذه البيانات هي من ناحية أولى وفي المقام الأول تهدف إلى وصف مصادر المعلومات الإلكترونية، كما أنها من ناحية أخرى لا تأتي مستقلة كما هو الحال مع تسجيلات الفهرسة في أحدث صيغها غير التقليدية كصيغة مارك، كما لا تأتي مقروءة للمستفيد حتى وإن جاءت مرافقة للمصدر كما هو الحال مع تسجيلات الفهرسة أثناء النشر المألوفة لنا، بل عادة ما ترافق المصدر ولكن غالباً ما تأتي بشكل خفي عن المستفيد وفي ملف فرعي آخر

(١) محمود عبد الكريم الجندي، المصدر السابق .

غير الملف الأصلي الذي يمثل المصدر، ومع ذلك لا أزال غير مقتنع قناعة كاملة بدقة دلالة هذه المصطلحات ومدى قبولها والارتياح لتداولها، لذلك وقياساً على مقولة " قليل شائع خير من كثير مهجور " استقر الرأي عندي على أكثر البدائل شيوعاً وتداولاً بين المختصين، أعني بذلك مصطلح "ما وراء البيانات"، وعليه كان هذا هو المصطلح العربي المتداول في هذا الكتاب مقابلاً للمصطلح الإنجليزي "metadata"، ولعل الأيام المقبلة تتيح لنا فرصة تلتقي فيها آراء الباحثين وجهوهم عند مصطلح من بين تلك المصطلحات المقترحة أو مصطلح جديد تتوافر فيه مقومات القبول لدى جميع المهتمين.

عن المصطلحات الأخرى :

وإذا كان هذا هو واقع الحال بالنسبة للمصطلح الأساسي الذي يعبر عن موضوع الكتاب وبؤرة اهتمامه ، فقد ارتبط بالكتاب من جهة أخرى أمر مثير وددت الإشارة إليه في هذا السياق ، ألا وهو كثرة المصطلحات الجديدة ، ليست على مستوى الساحة العربية فحسب ، وإنما على مستوى المجال عامة وتنظيم المعلومات خاصة ، ولا شك أن ترجمة هذه المصطلحات قد مثلت فيَّ تحدياً في كثير من الأحيان ، بل تطلبت ترجمتها وقتاً وجهداً إضافيين كادا يصرفانني في بعض الأحيان عن الترجمة برمتها ، وخاصة أن الهواجس السلبية عن مردود الترجمة ما تفتأ تنقطع عن مطاردة كل باحث يطمح أن يلج باب الترجمة في أوساطنا الأكاديمية العربية ، حيث تعامل الترجمات في بعض المجتمعات الأكاديمية العربية على أنها من الأعمال الثانوية أو غير الأصيلة والتي لا ترقى لأن تكون جهدًا علميًا؛ ومن

ثم لا تحتسب ضمن متطلبات الترقي في الدرجات الوظيفية . على أية حال فقد بذلت قصارى جهدي في ترجمة هذه المصطلحات، وعمدت إلى ذكر الترجمة العربية للمصطلح مصحوبة بالمصطلح في صيغته الإنجليزية هذا من ناحية، وفضلاً عن ترجمة المصطلحات الجديدة، شاب كثير من هذه المصطلحات والمفاهيم عدم وضوح فرأيت في مواضع كثيرة وما لم يكن من الممكن فهم هذا المعنى من السياق الذي وردت فيه إلى الاجتهاد في التعريف بهذه المصطلحات في هوامش صفحات الكتاب بحيث بدت كثير من الصفحات وقد تضمنت أكثر من حاشية ما بين شرح أو توضيح أو تعليق . ولإثبات مسؤوليتي عن تلك التعليقات أو التوضيحات وتمييزها عن الإرجاعات الببليوجرافية التي جاءت في متن الكتاب الأصلي فقد أتبعتها كلمة [المترجم] ، بل دفعتني كثرة هذه التعليقات إلى أن أشير في صفحة العنوان إلى أن جهدي فيه لم يقتصر على الترجمة بل تعداها إلى التعليق، وهكذا بدت مسؤوليتي فيه الترجمة والتعليق .

وتبقى كلمة :

وأخيراً فلم يتبق إلا واجب الشكر والتقدير لأستاذي الفاضل الأستاذ الدكتور محمد فتحي عبد الهادي الذي كان له فضل التنبيه على أهمية ترجمة هذا الكتاب، وكذا الزملاء الأفاضل الذين لم يكن لتشجيعهم وعونهم حدود، وخاصة الدكتور محمود الجندي، الذي تفضل بقراءة مسودة الترجمة وأبدى كثيراً من التعليقات القيمة، والزميلين الفاضلين الدكتور مبارك بن سعد سليمان والدكتور سعيد العسيري على تشجيعهما وحرصهما على أن ترى هذه الترجمة وغيرها من بحوث النور، كما أن الشكر موصول ومقدر

لمكتبة الملك فهد الوطنية ممثلة في إدارة النشر ، على رعايتها برنامج نشر الإنتاج الفكري العربي في مجال المكتبات والمعلومات عامة وموافقتها على نشر الكتاب ضمن هذا البرنامج، وأخيراً كل الشكر والمحبة لأفراد أسرتي الصغيرة على تحملهم عناء انشغالي لنحو خمسة عشر شهراً قضيتها في ترجمة هذا الكتاب، كما أسأل المولى تبارك وتعالى أن ينفع بهذا الكتاب، وأن يكون جهدنا فيه خالصاً لوجهه الكريم، وآخر دعوانا أن الحمد لله رب العالمين.

المترجم

الرياض في غرة جمادى الأولى من عام ١٤٢٧هـ.

تصدير المؤلفة

قصد بهذا الكتاب أن يكون مدخلاً للمكتبيين وغيرهم من العاملين في بيئة المكتبات إلى ما وراء البيانات، كما أتطلع إلى أن يفيد منه مدرسو علم المكتبات والمعلومات وطلابه في هذا الصدد.

وقد رأيت أن يخصص الجزء الأول من الكتاب لمعالجة المفاهيم والقضايا العامة المرتبطة بجميع خطط ما وراء البيانات، وأن يخصص الجزء الثاني منه للتعريف بعدد من معايير أو خطط ما وراء البيانات، وقد حاولت أن أتناول هذه الخطط من حيث أنواع مصادر المعلومات التي تسعى إلى وصفها، وتطبيقاتها، والكيفية التي توصف بها، ثم بيان أقسامها وعناصرها الرئيسة، والتركيبات التي ترد بها، وقواعد المحتوى التي تطبقها، وأخيراً عرض التطبيقات التي تستخدمها. وأود أن أشير إلى أنني ما قصدت بهذا الكتاب أن يكون دليلاً إرشادياً لكيفية فهرسة مصادر المعلومات أو وصفها اعتماداً على أي من خطط ما وراء البيانات التي تضمنها، ولكن حسبه أن يقدم فكرة عامة عن مدى تفاوت خطط ما وراء البيانات في توجهاتها لوصف مصادر المعلومات من جهة، ومدى تقاربها حول بعض الجوانب من جهة أخرى، وكيف أنها طورت لتحقق أهدافاً متباينة.

كما أود الإشارة كذلك إلى أن هذا الكتاب لا يتضمن حصراً شاملاً لجميع خطط ما وراء البيانات، فتلك مهمة يصعب تحقيقها في هذا السياق، كما أنه ليس من السهولة بمكان وضع معايير محددة يتم على أساسها اختيار ما ينبغي أن يتضمنه واستبعاد ما لا ينبغي أن يتضمنه، ولكني حاولت بشكل

عـام أن أضـمنه الخطـط المتداولـة فـي عـالم المكتبات أو في أي بيئة أكاديمية مشابهة، وها أنذا أشعر بتقصير لعدم تناولي جميع الخطط المعنية بمعلومات المـتاحف، أو بمعاييـر الوسـائط المـتعددة، أو بغيـرها مـن الخطـط المعنـية بوصـف المجموعات المتخصـصة جداً، واكتفائي بمعالجة بعض الخطط في مجـالات محددة، ولـذا فإنـي أهيب بكل من له اهتمام مباشر بخطط ما وراء بـيانات معيـنة لـم يغطهـا هـذا الكـتاب، أن يـستخدم بعـض الأدلـة الإرشـادية المعنـية بمـا وراء البـيانات والتـي تـضمنتها قائمـة القـراءات الملحقـة بالفصل الأول مـن الكـتاب، ففيها مـا يحيل إلـى المـزيد مـن المعلـومات عن كثير من خطط ما راء البيانات الأخرى.

وقد رأيت في سياق حديثي عن أي من خطط ما وراء البيانات أن تكتب أسـماء عناصرها وخاصـة المأخوذة من خطط ما وراء بيانات معرّفة بصيغة كـتابة الحـرف الأول من الكلمة بشكله الكبير، هكذا: Title، Online Provider، Technical Infrastructure of Complex Object. وعـندما تكـون الخطـة معـرفة فـي صيغة لغة الترميز المعيارية العامة SGML، أو كتعريف لنوع الوثيقة في صـيغة لغـة الترميز الموسعة XML DTD، أو كمخطط محدد Schema، يشار إلـى عنصر ما وراء البيانات فيها بأسماء تيجانها المحصورة بين القوسين الـزاويين، هكـذا:< title > ، أو < classDecl >، أو < taxonomy >. أمـا قيم عناصـر مـا وراء البيانات فقد عمدت إلى تسجيلها بين علامتـي اقتباس، هكذا كمـا فـي الجملـة التالـية: إن قـيمة نـوع التـسجيلة قد تكـون "عمـلاً work" أو "صـورة image". كـذلك ذكـرت [أسـماء] الخـصائص وقـيمها بـين علامتـي اقتباس، هكذا كما في الجملة التالية: قيمة الخاصية "id" هي **"١٢٣٤٥"**. وعند

وصف المعلومات التي ترد في سياق خطط ما وراء البيانات نفسها، رأيت كذلك أن تكتب أسماء الخصائص التي توصف عن طريقها عناصر ما وراء البيانات بين علامتي اقتباس كذلك، هكذا كما في الجملة التالية: يوصف كل عنصر بيانات من حيث الخصائص التالية: "الاسم name"، و"الوسيمة label"، و"التعريف definition"، و"المصطلح المقيد vocabulary".

تتوجه المؤلفة بالشكر لإلين بورسيل Ellen Purcell، بمركز فلوريدا لأتمتة المعلومات على مساعدتها في إخراج الأشكال التي تضمنها الكتاب، كما أسجل تقديري الكبير لجميع الزملاء الذين تفضلوا بمراجعة مسودة الكتاب أو جزء منها، وهم: مورث باكا Murth Baca من معهد جيتي للبحوث، وكارين كويل Karen Coyle من مكتبة كاليفورنيا الرقمية، وميشيل فوكس Michael Fox من الجمعية التاريخية لولاية مينيسوتا، وروبيكا جوينثر Robecca Guenther من مكتبة الكونجرس، واستيفاني هاس Stephanie Haas من مكتبة جامعة فلوريدا، وجيرترورود كو Gertrude Koh من جامعة الدومينيكان، فقد كان لملحوظات هؤلاء جميعاً أثرها الكبير في تحسين النص وتدارك كثير من الأخطاء.

وأخيراً أتوجه بالشكر لتشارلز هاسبندس Charles Husbands من جامعة هارفارد الذي كان له فضل توجيهي أثناء إعداد الكتاب إلى معرفة المزيد أولاً عن فهرسة المكتبات حتى أتجنب الوقوع في أخطاء لا أعذر على سوء فهمها.

الجزء الأول
المبادئ الأساسية والممارسة

الفصل الأول
أساسيات ما وراء البيانات

ما وراء البيانات ، ما هي؟

مع أن "ما وراء البيانات Metadata " موضوع يحظى باهتمام كبير في علم المكتبات والمعلومات ، إلا أن هذا المصطلح ينتمي في الأصل إلى مجال علم الحاسب حيث تستخدم السابقة " meta " بمعنى " عن about " ، وقياساً عليه يكون معنى كلمة Metalanguage لغة عن لغات أو لغة تصف لغات أخرى ، وبناءً على ذلك يصبح معنى مصطلح " metadata" " بيانات عن البيانات " أو " البيانات التي تصف بيانات أخرى". وقد ظهر أول استخدام لهذه الكلمة بمعنى بيانات عن البيانات، في الطبعة الأولى التي نشرت في عام ١٩٨٨ من الموجز الإرشادي لدليل ناسا لتبادل الصيغ NASA's Directory Interchange format Manual .(١)

ومما يلفت الانتباه أن أول من وضع مصطلح " METADATA " ـ الذي يكتب على هذا النحو بالحروف الكبيرة ، أو هكذا " Metadata " ـ أي بالحرف الأول الكبير فقط ، هو جاك إي مايرز Jack E. Myers وكان ذلك في أواخر الستينات من القرن العشرين الميلادي . وفي عام ١٩٨٦م

(١) E. Paul Shelley and B. David Johnson, "Metadata: Concepts and Models," in Proceedings of the Third National Conference on the Management of Geosciences Information and Data, organized by the Australian Mineral Foundation, Adelaide, Australia, ١٨-٢٠ July ١٩٩٥،٤,١-٥, available at

http://www.ainet.com.au/web٪٢٠pubs/Papers/AMF٩٥/Shelley&Johnson.html. Accessed ١ June ٢٠٠٢.

استخدام المصطلح علامة تجارية لشركة " Metadata company " ، إحدى الشركات التي تقدم برمجيات وخدمات متخصصة في الطب والرعاية الصحية ، وبصفتها علامة تجارية لهذه الشركة ، فإن كلمة "metadata " عندما تكتب كلمة واحدة على هذا النحو ، فهي تشير إلى المنتجات الحالية والمستقبلية لهذه الشركة ، أما المعاني الأكثر عمومية لهذه الكلمة فينبغي أن تتمثل بكتابتها إما في مقطعين يفصل بينهما بمسافة ، هكذا : "meta data" ، أو في مقطعين يفصل بينهما بشرطة قصيرة هكذا : "meta-data" ، ولهذا السبب فإن ائتلاف التعلم الكوني التابع لمؤسسة نظم الإدارة التعليمية[1] يسمي مواصفته لما وراء البيانات هكذا : IMS Meta-data Specification "، أي بكتابة الكلمة في مقطعين يفصل بينهما بشرطة، أما غالبية المبادرات الأخرى لما وراء البيانات فلم يكن من الصعوبة التعامل معها بهذا القدر، بل إن بعضها اكتسب شعبية واسعة بدرجة دخل معها مصطلح " metadata " إلى نطاق المجال العام.

ومع أوائل التسعينات من القرن العشرين الماضي، أصبح مصطلح " metadata " يستخدم بمعنى المعلومات الضرورية التي تيسر على الإنسان الإفادة من ملفات الحاسب ، وخاصة عند تعامله مع مجموعات البيانات[2]

---

(١) IMS Global Learning Consortium, inc. مع ملاحظة أن المختصر IMS يشير إلى Instructional Management Systems [المترجم] .

(٢) مصطلح مجموعة البيانات datasets عندما يكتب بالإنجليزية هكذا : dataset ـ كلمة واحدة ، أو هكذا: data set ، أي في مقطعين ، من المصطلحات المتداولة في سياق استخدام الحاسبات في العلوم الاجتماعية ، ويقصد به مجموعة من تسجيلات البيانات ، وإن كان هناك استخدام آخر للمصطلح في سياق بعض البرامج الإحصائية مثل برنامج " SAS "

المتخصصة في العلوم والعلوم الاجتماعية وذات الاتصال ومجموعات البيانات الجغرافية المكانية[١]. ومن أولى المواصفات التي وسمت نفسها بأنها مواصفة ما وراء بيانات، "معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية للجنة الفيدرالية للبيانات الجغرافية[٢] الذي نشرت إصدارته الأولى في عام ١٩٩٤م، وقد تمثلت الأهداف المعلنة لهذا المعيار في "تحديد مدى توافر إحدى مجموعات البيانات المكانية ، ومدى ملاءمة هذه المجموعة للاستخدامات المقصودة ، ثم تحديد وسائل الوصول إليها ، بالإضافة إلى مدى إمكانية نقلها بشكل آمن"[٣] .

ومع زيادة انتشار الانترنت والوب ، بدأ مصطلح " metadata " يستخدم في سياق وصف كيانات المعلومات الرقمية المتاحة على الشبكة ، وبخاصة بعد أن تبين أن الملفات النصية التي يسهل على الإنسان فهمها - على عكس ما هو عليه الحال بالنسبة لمجموعات البيانات العددية[٤] - تحتاج هي الأخرى إلى ما وراء بيانات تساعد على البحث عنها أو إيجادها، أو على الأقل إدارتها وضبطها. ثم دخل المصطلح بعد ذلك في حدود عام

ليشير إلى التمثيل الداخلي للبيانات ، ومن ثم نجده يرد هكذا في سياق هذا البرنامج المشار إليه SAS dataset [المترجم].

(٣) geospatial [المترجم].

(١) Federal Geographic Data Committee's Content Standard for Digital Geospatial Metadata [المترجم].

(٢) Federal Geographic Data Committee, Content Standard for Digital Spatial Metadata (CSDSM), version ٢،١٩٩٨, available at http://www.fgdc.gov/metadata/contstan.html. Accessed ١ June ٢٠٠٢.

(٢) numeric dataset [المترجم] .

١٩٩٥م إلـى مجـال المصطلحات العملية المتداولة فـي عـالم المكتبات ، مع استحداث مجموعة عناصر ما وراء البيانات الأساسية لدبلن ، أو ما تعرف اختصاراً بدبلن كور ؛ حيث كان القائمون على تنظيم أول ورشة لدبلن كور أعضاء نشطين في اتحاد الشبكة العالمية (الوب) - التي كانت حينئذ منظمة وليدة معنية بتطوير شبكة الوب، التي كانت هي الأخرى في أطوارها الأولى آنذاك. ولما كان هذا هو واقع الحال ، كانت مبادرة دبلن كور الأولى بمثابة منطقة تفاعل بين مجتمع المكتبات من ناحية ، ومجتمع شبكة الوب من ناحية أخرى ، وكانت قادرة بذلك على إمداد مجتمع المكتبات بأفكار ومصطلحات جديدة.

وسرعان ما اكتشف المكتبيون أنهم كانوا يقومون منذ قرون مضت بإنشاء بيانات عن البيانات وذلك في سياق عملية الفهرسة التي يمارسونها ، ومع ذلك لم يحظ هذا المصطلح بالقدر المطلوب من الاستخدام المطرد بين المكتبيين ؛ حيث حرص بعضهم على استخدامه لوصف المصادر الرقمية وغير الرقمية على حد سواء ، بينما قيد البعض الآخر منهم استخدامه على وصف المصادر الإلكترونية دون غيرها ، وخير مثال على تبني هذا المعنى الأخير موقع الوب الخاص بالاتحاد الدولي لجمعيات المكتبات ومؤسساتها- إفلا[١] - الذي يعرف هذا المصطلح بقوله : " إن هذا المصطلح يشير إلى

(١) the International Federation of Library Association and Institutions (IFLA) [المترجم].

أيـة بـيانات تـساعد علـى التثـبت مـن هـوية المـصادر الإلكتـرونية الـشبكية، والتعريف بها، ووصفها، وتحديد أماكنها"(١).

ومـع أن هـذا التفسير المقـيد كـان أقـرب ربمـا إلـى المفهـوم الأصـلي للمصطلح فـي سياق علـم الحاسـب، فـإن النظـر إلى "ما وراء البيانات" على أنهـا بيانات وصـفية لجميع أنـواع مـصادر المعلـومات يعتبـر بلا شك أكثر إفـادة، حيث تـندرج المـصادر المطـبوعة فـي إطـار هذا المفهوم، على اعتبار أن كثيـراً من المجموعات عادة ما تتألف من كيانات رقمية وغير رقمية، بل إن المجموعة الرقمية منها غالباً ما يكون لها أصولها غير الرقمية ، إلا أنني لا أريد أن أدخل في دائرة ما يثار من جدل حول القول بأن تسجيلة الفهرسة التـي تـصف مجلة إلكترونية تعد أحد أشكال "ما وراء البيانات"، بينما لا تعد التسجيلة التي تصف النسخة المطبوعة من المجلة ذاتها غير ذلك ؛ أي ليست من قبيل ما وراء البيانات.

وثمـة زعم آخر يثيـره البعض مفـاده أن "ما وراء البيانات" نفسها يجب أن تكون فـي صيغة إلكتـرونية بغض النظـر عن طبيعة الكيان الذي تصفه، وهذا الزعم ليس له ما يبرره من الناحية العملية لأن وصـف المصادر غالباً مـا يـتم الآن بـشكل رقمـي، كمـا أن اختـزان نـاتج هذا الوصـف غالباً ما يكون كذلك في صيغ رقمية، ومع ذلك فإن هذا الزعم قد يفهم منه أن تسجيلة مارك هـي "مـا وراء بيانات"، بينما بطاقـة الفهرسـة التـي لـم تحول بعد إلى صيغة

(٢) International Association of Library Associations and Institutions, Digital Libraries: Metadata Resources, available at http://www.ifla.org/II/metadata.htm. Accessed ١ June ٢٠٠٢.

مـارك ليست كذلك، أي ليست "مـا وراء بيانات"، ولعل فـي ذلـك مـا يثيـر حفيظة البعض فـي الـوقت الـذي يصر فيه آخرون علـى أن هذا هو الفارق الواضـح بـين هـذين النمطـين مـن التـسجيلات . ومـن الجديـر بالملاحظـة أن الناشـرين كانوا أكثر اقتناعاً بهذا التصور العام من كثير من المكتبيين. ولعل مـا يـؤكد ذلـك تعـريف جمعية الناشـرين الأمـريكيين "لمـا وراء البـيانات"، والتـي تنص علـى أنهـا " معلومات تـصف المحتوى، وأفضل مثـال واقعي علـى ذلـك الفهرس البطاقي في المكتبة، أو المدخل في الفهرس المطبوع في شكل كتاب، أو المعلومة في الكشاف المتاح على الخط المباشر"[1] .

وهناك تـوجه مغاير يحـيط بهذا المصطلح يتمـثل في ما لوحظ من تبني اتحـاد الوب العالمي[2] للتعريف الأكثر تقييداً من بين جميع التعريفات، وهو أن "مـا وراء البيانات هـي معلـومات مفهومة آلياً للوب[3] "[4] . وكون أن "مـا وراء البـيانات" ينبغـي أن تكون مفهومة آلياً يأتـي مناقضاً تماماً للحاجة الأصـلية التـي يعبر عنـها العلماء في مجالي العلوم والعلوم الاجتماعية الذين يـتعاملون مـع مجمـوعة البـيانات العدديـة، والـذين ينظـرون إلـى مـا وراء

(١) MICI Metadata Information Clearinghouse (Interactive) (home page). Accessed ١ June ٢٠٠٢. Available at http://www.metadatainformation.org/ superseded by the Semantic Web activity with a home page at http://www.w٣.org/٢٠٠١/sw/.

(١) W٣C [المترجم] .

(٢) machine – understandable [المترجم] .

(٣) Metadata and Resource Description (home page of the W٣C Metadata Activity). Accessed ١ June ٢٠٠٢. Available at http://www.w٣.org/Metadata . The Metadata activity was superseded by the Semantic Web activity with a home page at http://www.w٣.org/٢٠٠١/sw/.

البيانات كوسيلة لجعل بيانات حاسباتهم مفهومة للبشر[1]. وهذا التعريف يعبر أيضاً عن مطلب أكثر دهاءً يستنبط من القول بأن "ما وراء البيانات معلومات لشبكة الوب"، ساحباً بذلك الأهلية ليس فقط من الفهارس البطاقية والفهارس المطبوعة التي يتيحها الناشرون ، وإنما يسحبها كذلك من وصف المصادر التي يمكن الوصول إليها عن طريق جميع بروتوكولات الإنترنت الأخرى وغيرها من جميع النظم الآلية الأخرى غير المعتمدة على الوب.

إلى هنا يجب أن يكون واضحًا وبما لا لبس فيه ، أنْ ليس هنالك تفسير جامع مانع، أو صحيح أو غير صحيح "لما وراء البيانات"، ولكن ينبغي على من يتعامل مع هذا المصطلح أن يعلم أن مدلوله قد يختلف باختلاف المجتمع الذي يستخدمه، والسياق الذي يستخدم فيه، وفي هذا النص الذي بين يدي القارئ استخدمت تعريفاً غير مقيد إلى حد ما، فما وراء البيانات في هذا السياق يقصد بها معلومات مهيكلة[2] أو منظمة في صيغة معينة عن أحد مصادر المعلومات الذي يصدر في أي وسيط مادي أو في أي صيغة كانت، ومن الملاحظ أن هذا التعريف لا يتعرض لطبيعة هذه المعلومات المهيكلة، بمعنى هل هي إلكترونية أم غير إلكترونية، ولا إلى طبيعة المصادر الموصوفة، بمعنى هل هي إلكترونية أو متاحة عبر الشبكات[3] أو عن طريق الوب[4]، كما أنه لا يهتم كذلك بالطرف الذي توجه إليه المعلومات هذه ، بمعنى هل ما وراء البيانات موجهة لاستخدام الإنسان أم

(٤) human-understandable [المترجم] .
(١) structured [المترجم] .
(٢) network-accessible [المترجم].
(٣) web-accessible [المترجم] .

الآلة؛ لكن ما ينبغي أن يكون واضحاً أن هذا التعريف يضع شرطين يجب توافرهما لهذه المعلومات لتكتسب صفة " ما وراء بيانات "، أولهما أن هذه المعلومات يجب أن تكون منظمة أو موضوعة في صيغة محددة أو تركيبة بنائية معينة، أي لا تكون مجرد مجموعة عناصر بيانات مجمعة أو ممثلة بشكل عشوائي ، بل يجب تسجيلها طبقاً لخطة ما وراء بيانات جيدة التوثيق، أما ثانيهما فيتمثل في أن "ما وراء البيانات" يجب أن تصف أحد مصادر المعلومات، مع تسليمنا بداهة بماهية مصدر المعلومات على النحو الذي سلمت به ورشة عمل دبلن كور عام ١٩٩٥م، وعرفته بأنه "كيان وثائقي، أو كيان تتوافر له مقومات الوثيقة"(١) .

وهنالك جانب آخر نود الإشارة إليه في هذا الصدد وهو أن "ما وراء البيانات" موضوع يحظى كذلك باهتمام مكثف في المجالات المتصلة بالأعمال التجارية والتصنيع والتجارة الإلكترونية ، حيث تستخدم مثلاً لوصف البراغي(٢) والعدد(٣) وأطقمها ، والمعاملات التي تتصل بذلك. ولا شك أن الوصف المنظم الضروري لضبط تصنيع تلك المواد وتخزينها وتسويقها هو بالتأكيد عبارة عن "ما وراء بيانات"، ومهما تعددت المبادرات في هذا الصدد، فكلها مهتمة بالتوسع في تعريف "ما وراء البيانات" وتبادلها ، غير أني رأيت أن أستبعد جميع هذه المبادرات عن

(٤) document-like object [المترجم] .

(٥) screws [المترجم] .

(١) widgets [المترجم] .

سياق هذا الكتاب ، ليس لعدم أهميتها ، وإنما لأنني حقاً لا أعرف شيئاً عن هذه الأشياء تماماً .

وبشكل عام يمكن القول إن من أكثر المناقشات التي يعول عليها فيما يتصل بما وراء البيانات ، هي أنها لا تهتم في مجملها بالمجال الذي تطبق فيه أو بالمصدر الذي يطبق عليه، ولكن بالهدف الذي يؤمل منها أن تنجزه، ولعل خير مثال على هذا التوجه ما جاء في أحد معاجم المصطلحات التي نشرها معهد جيتي للأبحاث(١)، حيث عرف تلك البيانات التي ترتبط إما بنظام معلومات معين أو بكيان معلومات محدد ، بهدف وصفه وإدارته، وإقرار متطلباته القانونية، ووظائفه الفنية ، ثم دعم مقومات استخدامه والإفادة منه وحفظه"(٢)، وقد نحى مكتب شبكات المكتبات والمعلومات البريطانية(٣) هذا المنحى بقوله: إن ما وراء البيانات تعني البيانات المهيكلة أو المنظمة عن المصادر الرقمية (وغير الرقمية)، بحيث يمكن استخدام هذه البيانات في دعم مجموعة كبيرة من العمليات، منها - على سبيل المثال -

(٢) Getty Research Institute [المترجم] .

(٣) Murtha Baca, ed., Introduction to Metadata: Pathways to Digital Information, version ٢,٠, available at http://www.getty.edu/research/institute/standards/intrometadata . Accessed ١ June ٢٠٠٢.

(٤) U.K. Office for Library and Information Networking [المترجم] .

وصف هذه المصادر واستكشافها، وإدارتها (بما في ذلك إدارة الحقوق)، هذا فضلاً عن حفظها حفظاً طويل الأمد"(١).

أنواع "ما وراء البيانات" :

أدى الاعتراف بالاستخدامات المتعددة لما وراء البيانات إلى وجود أرضية متسعة لها بحيث تتنوع بين ثلاث فئات أساسية هي : ما وراء البيانات الوصفية، و ما وراء البيانات الإدارية ، ما وراء البيانات البنائية أو التركيبية. ومن الملاحظ أن هذه الفئات تشير إلى مدى الاستخدامات الوظيفية لما وراء البيانات والغرض منها أكثر من إشارتها إلى القيم الكبيرة التي تتسم بها عناصر ما وراء البيانات ؛ ذلك أن جميع "ما وراء البيانات" تعد- من حيث التعريف - واصفة لشيء ما.

وبالنسبة لما وراء البيانات الوصفية(٢) فهي التي تساند عمليات استكشاف مصادر المعلومات (أي كيفية إيجادها)، والتعريف بها (أي كيفية تمييز المصدر عن غيره من المصادر الأخرى المشابهة له) ، واختيارها (أي كيفية تحديد أن مصدراً ما يلبي احتياجات معينة ، مثال ذلك نسخة على أسطوانة دي في دي DVD من أحد أشرطة الفيديو)، وفضلاً عن ذلك يمكن استخدام هذه الفئة من ما وراء البيانات - أي الوصفية - كذلك في عمليات التجميع(٣) (أي تجميع كل نسخ العمل الواحد بعضها مع بعض)،

---

(١) Metadata (page of the UKOLN website). Accessed I June ٢٠٠٢. Available at http://www.ukoln.ac.uk/metadata/

(٢) descriptive metadata [المترجم].

(٣) Collocation [المترجم] .

والاقتناء (أي الحصول على نسخة من مصدر المعلومات أو الوصول إليها) . وإذا ما نظرنا إلى الفهرسة التقليدية للمكتبات على أنها ما وراء بيانات فإنها تدخل أساساً في إطار فئة ما وراء البيانات الوصفية ، وهو ما يسري كذلك على خطتي ما وراء البيانات لكل من دبلن كور ، وجمعية المصادر المرئية(١).

وهناك وظائف أخرى تدخل في نطاق المعنى الواسع لما وراء البيانات الوصفية، من أهمها وظائف التقييم ، والربط ، ومدى القابلية للاستخدام. مع ملاحظة أن التقييم قد يكون تقييماً سردياً أو ذاتياً كما يحدث عند القيام بعرض أحد الكتب أو الأفلام ، أو أن هذا التقييم قد يعبر عنه بشكل رسمي من خلال أسلوب تقديرات المحتوى(٢) الذي تستخدم بموجبه إحدى خطط التقدير التي ترعاها بعض الهيئات ، كخطة تقدير الأفلام المتحركة(٣) (التي تعتمد مستويات التقدير التالية : G, PG, PG-٣١) التي ترعاها جمعية الأفلام المتحركة الأمريكية(٤)، ووسيمات جمعية تقييم محتوى الإنترنت(٥) .

أما عملية الربط(٦) فهي التعبير عن العلاقات بين المصدر الموصوف من جهة وأي مصدر آخر أو مجموعة من المصادر الأخرى من جهة أخرى ، مع ملاحظة أنْ ليس هناك حد معين لعدد العلاقات المحتملة ؛ فقد يُربط أحد الكتب الموصوفة مثلاً بطبعاته السابقة أو اللاحقة وبترجماته

(١) VRA (Visual Resources Association) [المترجم] .
(٢) content ratings [المترجم] .
(٣) the movie rating scheme [المترجم] .
(٤) the Motion Picture association of America [المترجم] .
(٥) the Internet Content Rating association labels [المترجم] .
(٦) linkage [المترجم] .

وبإصداراته الأخرى، بل بكتب أخرى للمؤلف نفسه، أو في الموضوع ذاته، كذلك فإن مقالة من مقالات الدوريات قد تربط بالدورية التي نشرت فيها، أو بالمقالات الأخرى في العدد نفسه، وبالأعمال التي استشهدت بها، كما أن الكيان الرقمي قد يربط بكيانات أخرى تعبر عن المحتوى نفسه ولكن بصيغ أو في وسائط أخرى ، مثال ذلك أحد المباني، والصورة الضوئية لهذا المبنى، والصورة المأخوذة بالماسح الضوئي في صيغة تيف TIFF عن هذه الصورة الضوئية ، وملف العرض بصيغة ملف الصور، أو ما تعرف بصيغة بي إس دي PSD الخاصة ببرنامج الفوتوشوب المشتق من صيغة تيف تلك، والتحسينات التي يجريها أحد الفنانين على هذه الصيغة الأخيرة (أي صيغة PSD)، وصفحة الإنترنت المصممة بلغة تهيئة النص الفائق HTML التي تتضمن الصورة التي صورها ذلك الفنان ، وموقع الإنترنت الذي تعد هذه الصفحة أحد مكوناته ، فكل هذه الصيغ يمكن ربطها بعضها ببعض بأساليب يمكن تحديدها والتعبير عنها في سياق ما وراء البيانات.

ومن الملاحظ أن للقدرة على التعبير عن العلاقات الموجودة بين الكيانات الرقمية أهمية خاصة ، نتيجة لإمكانية استخدام أحد الروابط الفعالة actionable أو الفائقة hyperlink للربط فيما بينها من ناحية ، ونتيجة لتكاثر صيغ مصادر المعلومات وإصداراتها المتنوعة في البيئة الرقمية من ناحية أخرى، ولهذا السبب فقد أشارت لجنة المجلس الوطني للأبحاث National Research Council Committee في إحدى استراتيجيات تقنية المعلومات لمكتبة الكونجرس إلى أن "عدم صلابة plasticity " الكيانات الرقمية سوف تدعو إلى

الاهتمام بقضايا العلاقات بشكل يزيد عما كان يولى من اهتمام للمصنوعات الفنية المادية physical artifacts "(١).

وتتطلب الكيانات الرقمية كذلك مزيدًا من "ما وراء بيانات" تتصل بمدى القابلية للاستخدام usability ؛ لأنه وإن كان أغلب الناس يمكنهم استخدام إحدى الوثائق المطبوعة دون مساعدة ، إلا أن محتوى هذه الوثيقة نفسها سيحتاج لبعض التفسير إذا كانت بصيغة لغة الترميز المعيارية العامة SGML ؛ ولهذا السبب تسمح قواعد الفهرسة في المكتبات بتسجيل معلومات عالية المستوى عن مدى القابلية للاستخدام ، مثل تحديد الأجهزة والبرمجيات اللازمة لاستخدام المصادر الإلكترونية. كذلك تسمح مبادرة ترميز النص Text Encoding Initiative المعنية بترميز النصوص في صيغة لغة الترميز المعيارية العامة ، ليس فقط لغرض قراءتها ولكن لأغراض تحليلها لغوياً وأدبياً ، بتسجيل وصف تفصيلي ودقيق عن كيفية ترميز المحتوى النصي ؛ من هنا تعين على خطط "ما وراء البيانات" المعنية بتوثيق مجموعات البيانات datasets أن تتضمن معلومات شاملة عن المنطق الذي اختيرت على أساسه عناصر بيانات هذه الخطط والبنية التي توضع فيها .

(١) National Research Council, Committee on an Information Technology Strategy for the Library of Congress, LC٢١: A Digital Strategy for the Library of Congress (Washington, D.C.: National Academy Press, ٢٠٠١), available at http://books.nap.edu/catalog/٩٩٤٠.htrnl?onpinewsdoc٠٧٢٦٠٠ . Accessed I June ٢٠٠٢.

أما " ما وراء البيانات الإدارية administrative metadata " فهي عبارة عن معلومات تهدف إلى تيسير إدارة المصادر، ومن ثم تتضمن هذه الفئة معلومات عن الوقت الذي أنشئ فيه الكيان الرقمي ، والكيفية التي أنشئ بها ، والشخص المسؤول عن ضبط إتاحته أو الوصول إليه ، وطرق حفظه ، وبالإضافة إلى بيان عمليات المعالجة التي أجريت عليه ، والقيود المفروضة على إتاحته أو الوصول إليه ، وعلى الإفادة منه أو استخدامه.

ومن الواضح أن الفارق بين " ما وراء البيانات الوصفية " و " ما وراء البيانات الإدارية " ليس فارقاً قاطعاً ، وإنما يعتمد غالباً على رؤية المستفيد من "ما وراء البيانات". حيث ينظر إلى رقم القيد accession number - على سبيل المثال - على أنه أحد أشكال ما وراء البيانات الإدارية عند استخدامه في سياق إجراءات الحصول على المواد التي تقوم بها إحدى الجهات ، ومع ذلك ، فإن هذا الرقم - وطالما أنه أحد المعرفات الفريدة للمادة - فيمكن استخدامه أيضاً أداة وصفية لتعريف تلك المواد ، وبالمثل يمكن للباحث في حال بحثه فقط عن المواد المتاحة على نحو مباشر فقط دون غيرها من المواد ، استخدام قيود الإتاحة أو الوصول ؛ فتلك القيود التي هي من قبيل ما وراء البيانات الإدارية ، ساعدت إذن في وظيفة الاختيار(١). ومع هذا الغموض، فإن للفارق بينهما أهميته على مستوى الممارسة ، لأن كلاً من ما وراء البيانات الإدارية و الوصفية غالباً ما تُعرّفان في خطط مختلفة ، وتستخدم في نظم مختلفة من قبل أناس مختلفين ، مثال ذلك أن ما وراء البيانات الوصفية عادة ما تتوافر في نظم البحث

(١) من الملاحظ أن وظيفة الاختيار تدخل ضمن مهام ما وراء البيانات الوصفية [المترجم].

المتاحة للجميع ، بينما يقتصر الاطلاع على "ما وراء البيانات" الإدارية على المسؤولين عن إدارة مصادر البيانات.

ويمكن تقسيم "ما وراء البيانات" الإدارية إلى فئات فرعية يصعب حصرها وإن كان من أهمها : "ما وراء بيانات إدارة الحقوق rights management metadata "، و"ما وراء بيانات الحفظ preservation metadata " ، و"ما وراء البيانات الفنية technical metadata ". وتبدو وظيفة "ما وراء بيانات إدارة الحقوق" و"ما وراء بيانات الحفظ" واضحة من تسميتهما ، أما "ما وراء البيانات الفنية" فتتركز وظيفتها على توثيق خصائص الملفات الرقمية ، وغالباً ما تتم على مستوى عال من التفصيل؛ كأن يوضح مثلاً الكيفية التي يجزأ بها أحد الملفات المكتوب في صيغة تيف TIFF ، بمعنى هل هو مجزأ إلى مربعات أم إلى شرائح . و عليه فإن ما وراء البيانات الفنية تعد أحد المكونات الهامة لما وراء بيانات الحفظ ، حيث لا يمكن إعادة بناء تلك الملفات أو تحويلها إلى أشكال أو صيغ أخرى إلا بعد معرفة الخصائص المادية التفصيلية عن تلك الملف.

أما وراء البيانات التركيبية أو البنائية structural metadata " فيمكن النظر إليها - ومن باب التشبيه - وكأنها المادة الصمغية التي تعمل على تماسك الكيانات الرقمية بعضها مع بعض . فالكتاب مثلاً قد يحتوي على عدة فصول، وكل فصل منها يتألف من عدة صفحات ، ومن ثم فإن كل صفحة من هذه الصفحات يمثلها ملف رقمي منفصل ؛ وهنا تكمن وظيفة " ما وراء البيانات التركيبية " في تسجيل الروابط بين الملفات والصفحات المادية من ناحية ، وبين الصفحات والفصول من ناحية ثانية ، وبين

الفصول والكتاب كله من ناحية ثالثة. وبهذا المنطق فإن برمجيات العرض تستخدم " ما وراء البيانات التركيبية " لعرض قوائم المحتويات ، وتوظيفها كذلك في عمليات الانتقال المباشرة إلى فصل معين، أو تقليب الصفحات للأمام أو للخلف . كما أن " ما وراء البيانات التركيبية" تربط مكونات الوسائط المتعددة بعضها ببعض ، كان تعمل على ربط الصوت مع النص المكتوب ، وضبط تزامن صوت المعلق مع النص المكتوب، كما تقوم كذلك بتوثيق طريقة ترتيب عناصر البيانات والصيغ التي ينبغي أن تظهر بها في مجموعات البيانات العددية أو الإحصائية ، كبيانات الإحصاء السكاني مثلاً. ونظراً لأن الهدف الأساسي لما وراء البيانات التركيبية يتمثل في تمكين المستفيد من استخدام الكيانات entity [١]، فإنها تشترك مع " ما وراء بيانات قابلية الاستخدام الوصفية descriptive usability metadata " في بعض الجوانب ، وإن كان ثمة فارق بينهما فإنه يتمثل في أن "ما وراء بيانات مدى قابلية الاستخدام " موجهة أساساً لاستخدام البشر ، بينما تستخدم " ما وراء البيانات التركيبية " عموماً في عمليات المعالجة التي تقوم بها الحاسبات.

خطط "ما وراء البيانات" :

من الممكن التفكير في عناصر "ما وراء البيانات" كل على حدة. فإذا افترضنا- جدلاً - أن لدينا رفاً يحتوي على مجموعة من الكتب، وقد طبع على كعب كل كتاب منها عنوانه، فإن هذه العناوين هي شكل من أشكال "ما وراء البيانات"، وهي بالتأكيد إحدى طرق التعرف إلى هذه الأعمال ،بل

(١) غالباً ما تشير كلمة entity إلى أحد مكونات الكيان الرقمي [المترجم] .

تكون بديلاً عن سحب كل كتاب من الرف وقراءة عنوانه من صفحة العنوان. (في واقع الأمر لم تبدأ طباعة العناوين على كعوب الكتب إلا في النصف الأول من القرن الثامن عشر، وقبل ذلك الحين كان يتم استخدام "ما وراء بيانات" خارجية في شكل قوائم تحدد مواضع الكتب على الأرفف)(١). ولكننا على أية حال ننظر للموضوع بشكل عام من زاوية " خطط ما وراء البيانات" التي هي عبارة عن مجموعة من عناصر ما وراء البيانات ، مع قواعد استخدامها ، وقد قصد بهذه المجموعة أن تحقق هدفاً معيناً. وفي نطاق هذا التعريف العام ثمة مصطلحان غالباً ما يستخدمان على نحو تبادلي ، هما : scheme ، و schema (٢)، ونظراً لأن المصطلح الأخير يعنى أمرين آخرين ، يتصل أولهما بتقنية قواعد البيانات الآلية بمعنى التنظيم أو البناء الرسمي لقاعدة البيانات ، بينما يتصل الآخر بلغة الترميز الموسعة XML ويأخذ معنى خاصاً في سياق هذه اللغة، فإننا نفضل في سياق هذا الكتاب استخدام المصطلح الأول ، أي scheme .

ونعرف في الفقرات التالية بإيجاز ببعض الخطط الوصفية التي تتواتر الإشارة إليها بكثرة في سياق هذا الجزء الأول من الكتاب، أما التفاصيل الكاملة لخطط محددة فسيكون محلها في سياق الجزء الثاني من الكتاب .

- مجموعة دبلن لعناصر ما وراء البيانات الأساسية، المعروفة اختصاراً بدبلن كور DUBLIN CORE ، هي مجموعة بسيطة من

---

(١) Henry Petroski, The Book on the Bookshelf (New York: Alfred A. Knopf, ١٩٩٩).

(٢) كلاهما يعني بالعربية خطة أو مخطط [المترجم].

خمسة عشر عنصراً ، الغرض منها أن تكون عامة التطبيق على جميع أنواع مصادر المعلومات ، وقد طورت مبادرة دبلن كور لما وراء بيانات Dublin Core Metadata Initiative هذه المجموعة من العناصر الأساسية وأقرها كل من معهد المعايير الوطنية الأمريكية ANSI ، والمنظمة الوطنية لمعايير المعلومات NISO ، وتعرف الآن بمعيار ANSI/NISO Z٣٩,٨٥. ويمكن الاطلاع على الوصف المرجعي لمجموعة العناصر على الموقع التالي : http://dublincore.org/documents/dces/ ، كما أن عنوان الصفحة الرئيسية لمبادرة دبلن كور لما وراء بيانات هو http://dublincore.org. .

- الفئات الأساسية لجمعية المصادر المرئية Visual Resources Association core Categories (VRA Core) ، طورت هذه المجموعة أساساً لوصف المواد الموجودة ضمن مجموعات مصادر مرئية والتي تشتمل عادة على بدائل لأعمال أصلية ذات صبغة فنية ومعمارية (ومن أمثلة هذه البدائل: الصور الضوئية ، وشرائح العرض أو الصور الرقمية، أو كلاهما معًا). وهناك تعريف بهذه المجموعة من العناصر على الموقع التالي: http://www.vraweb.org/vracore٣.htm ، وعلى الصفحة الرئيسية لجمعية المصادر المرئية http://www.vraweb.org/.

- الوصف الأرشيفي المرمز Encoded Archival Description (EAD) ، أعد الوصف الأرشيفي المرمز كطريقة لتمثيل معينات البحث الأرشيفية في صيغة إلكترونية ، ومعينات البحث هي شكل من أشكال الوصف الأرشيفي الذي يبدأ عادة بمعلومات عامة عن المجموعة كلها، ثم يقدم وصفاً تفصيلياً عن مكوناتها. وعنوان الموقع الرسمي للوصف الأرشيفي المرمز هو: http://www.loc.gov/ead/ead.html.

- القواعد الأنجلو أمريكية للفهرسة ـ الطبعة الثانية / الفهرسة المقروءة آليًا (مارك) AACR٢/MARC ، لا تعد هذه القواعد بالضبط خطة ما وراء بيانات على النحو الذي اعتبرت به مجموعات عناصر ما وراء البيانات السابقة بمثابة خطط ما وراء بيانات ، غير أن كلاً من مجموعة القواعد التي تتضمنها ووصف الصيغ التي ترد بها

عناصر المعلومات في الفهرسة التقليدية في المكتبات يمكن أن تشكل مجتمعة وعلى نحو وظيفي خطة لما وراء بيانات. ويدخل في إطار هذه الفئة كل من التقنين الدولي للوصف الببليوجرافي ، والقواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية المنقحة (AACR٢R) ومارك ٢١ ، ويمكن الاطلاع على معلومات عن التقنين الدولي للوصف الببلوجرافي على موقع الاتحاد الدولي لجمعيات المكتبات ومعاهدها http://www.ifla.org/VI/٣/nd١/isbdlist.htm، أما قواعد الفهرسة الأنجلو أمريكية - الطبعة الثانية، التي أعدتها جمعية المكتبات الأمريكية، وجمعية المكتبات الكندية ، وجمعية المكتبات البريطانية فلا تتوافر إلا في إصدارات مطبوعة أو على أسطوانات مدمجة ، أما مارك٢١ والمواصفات ذات الصلة فتتاح على الموقع التالي:

http://lcweb.loc.gov/marc/

وثمة ثلاثة جوانب ترتبط بأي خطة من خطط "ما وراء البيانات" ، وهي: الجوانب الدلالية لعناصر البيانات أو ما تعرف بالدلالات semantics ، وقواعد المحتوى content rules، والتركيبة أو البنية التي ترد فيها عناصر البيانات Syntax . وتشير الدلالات إلى معاني عناصر ما وراء البيانات نفسها، وعادة ما تُعرّف خطة "ما وراء البيانات" عناصرها عن طريق إعطاء كل منها اسم وتعريف، كما يجب كذلك أن تبين الخطة بالنسبة لكل عنصر من عناصر بياناتها طبيعة هذا العنصر ، بمعنى هل هو عنصر إجباري أم اختياري أو إجباري عند الضرورة conditionally required (بمعنى أنه يكون إجبارياً عندما تدعو الحاجة إلى ذلك mandatory if

applicable) ، وأن تبين كذلك مدى تكرار هذا العنصر، بمعنى هل هو من العناصر التي يتكرر ذكرها في التسجيلة ، أم من العناصر التي لا تسجل إلا مرة واحدة ولا تكرر.

أما قواعد المحتوى فتحدد الكيفية التي يتم بها اختيار وتمثيل قيم عناصر "ما وراء البيانات" ، ففي الوقت الذي تعنى فيه دلالات خطة "ما وراء البيانات" بوضع تعريف لعنصر البيانات المسمى " المؤلف" - مثلاً - تحدد قواعد المحتوى المعلومات التي يتم بموجبها اعتبار أحد العناصر ليكون هو "المؤلف" (وهذه هي وظيفة الاختيار) ، ثم تبين القواعد كذلك الكيفية التي يجب أن يسجل بها اسم هذا العنصر الذي اعتبر مؤلفاً (وهذه هي وظيفة التمثيل representation)؛ فالقواعد الأنجلو أمريكية للفهرسة - على سبيل المثال - تنص على استخدام الشكل الأكثر شهرة لاسم الشخص، ثم تعطي عدداً من القواعد الفرعية لكيفية التثبت من أن هذا هو الشكل المشهور بالفعل، وقد تحدد مجموعة أخرى من قواعد المحتوى صيغة معيارية أو مقننة للاسم، كالاسم الأخير أو أول حرف من الاسم الأول أو الأوسط، بينما قد تطلب مجموعة أخرى من القواعد من منشئ "ما وراء البيانات" تقديم معرِّف فريد للمؤلف اعتماداً على أحد الملفات الاستنادية.

أما التركيبة بالنسبة لخطة ما وراء البيانات فتنصرف نحو كيفية ترميز العناصر في شكل مقروء آلياً ، ومن الناحية العملية يمكن القول إن نظم المعالجة المصممة لبحث ما وراء البيانات أو عرضها أو التعامل معها ، قد يكون لها صيغ تخزين داخلية مختلفة تماماً عن صيغة "ما وراء البيانات"، وبالتالي فإن التركيبة المحددة لخطة ما وراء البيانات تقدم صيغة تبادل

موحدة لتداول "ما وراء البيانات" بين الأطراف المعنية ، أكثر من وضعها قيوداً على كيفية تخزين البيانات في أي نظام محلي. ولهذا السبب فإن تركيبة خطة "ما وراء البيانات" يمكن تسميتها شكل اتصال communications format ، أو صيغة تبادل exchange format ، أو تركيبة نقل transport syntax ، أو تركيبة تراسل transmission syntax .

ومع أن الدلالات وقواعد المحتوى والتركيبة يمكن اعتبارها من الناحية المنطقية أشياء مستقلة ، إلا أنها ترتبط ببعضها ارتباطاً وثيقاً ، وعلى مستوى الممارسة فإن أي خطة ما وراء بيانات قد تشتمل على هذه المكونات ، مستقلة عن بعضها عن بعض ، أو قد تدمج بعضها مع بعض، أو قد يحذف بعضها وذلك في أي توليفة تراها مناسبة . حيث تقوم بعض خطط "ما وراء البيانات" بتحديد هذه المكونات في صيغة اللغة المعيارية الموحدة للترميز SGML ، أو في صيغة لغة الترميز الموسعة XML دون وضع حدود فاصلة بين الدلالات والتركيبة ، هذا في الوقت الذي أخفقت فيه خطط أخرى لما وراء البيانات في تحديد تركيبة معينة ، أو وضع قواعد لمطبقي هذه القواعد يتم بموجبها إتاحة الفرصة لهم لعمل تركيبات متعددة معتمدة ، كما أن هناك خططاً أخرى ليس لها قواعد محتوى خاصة بها ، وبعضها يحيل إلى قواعد محتوى خارجية ، أو أنها تكون مصممة بشكل يسمح باستخدام أي قواعد محتوى شريطة أن تكون مراعية لقواعد الخطة.

وغالباً لا تكون الجوانب الدلالية وقواعد المحتوى والتركيبة التي تحددها خطة "ما وراء البيانات" جامعة مانعة، وإنما تعرض على مَنْ يرغب في تطبيقها بدائل كثيرة ، وفي هذه الحالة لا تكون قواعد الخطة كافية وحدها

لضمان ثبات "ما وراء البيانات" سواء التي أنشأها أفراد مختلفون أو حتى التي أنشأها الشخص نفسه على فترات متباعدة. ومن المتداول بين مستخدمي خطط "ما وراء البيانات" اتباع إرشادات أكثر تقييداً من الخطة نفسها ، مع ملاحظة أن هذه القيود قد يقتصر استخدامها في مجال مشروع أو قسم معين ينشئ "ما وراء بيانات" ، أو يمكن تقاسمها بين مجتمع محلي ومجتمع دولي. وهنالك ما يعرف بالملفات التفصيلية أو البروفايلات Profiles (التي تعرف كذلك ببروفايلات التطبيق) وهي عبارة عن مواصفات طورت بشكل رسمي لتقييد حدود استخدام خطة "ما وراء البيانات" على مجتمع مستفيدين معين ، مع توضيح طبيعة هذا الاستخدام ، وأياً ما كانت طبيعة هذه القواعد في شكل إرشادات غير رسمية أو بروفايلات رسمية ، تظل هنالك حاجة لا محيد عنها لإيجاد مجموعة قواعد إضافية تلحق بخطط "ما وراء البيانات" عند نشرها.

وحتى في داخل عالم المكتبات ومؤسسات التراث الفكري الحضاري وعالم النشر وخدمات المعلومات المحدود ، يتم تبني عدد كبير من خطط "ما وراء البيانات" المتنوعة في محتواها وأشكالها . حيث رأت خطة دبلن كور - على سبيل المثال - أن تخصص لكل عنصر من عناصرها اسمًا ومعرفاً وتعريفاً وشرحاً، كما في هذا المثال :

Name: Title
Identifier: Title
Definition: A name given to the recourse
Comment: Typically, a Title will be a name by which the resource is formally known.

هذا بالإضـافة إلـى المقيدات أو المحددات qualifiers (أي الشروط التي تحدد معنـى أحد العناصـر ، أو تشير إلى كيفية تمثيل قيمته) والتي عادة ما تكون معرفة فـي وثيقة مستقلة فـي حالة عدم وجود قواعد محتوى ملزمة . وهـناك كـذلك عناصـر المقاربـة أو الـروابط mappings التـي أعـدت لتـربط عناصـر دبلن كور مع مجموعات أخرى من عناصـر "مـا وراء بيانات" ، لكنها وضعت مستقلة عن خطة دبلن كور نفسها.

وعلـى النقـيض مـن ذلـك ، فـإن الإصـدارة رقـم ٣,٠ مـن المجمـوعة الأساسية لجمعية المصادر المرئية VRA Core ، تحدد بالنسبة لكل عنصر بيانات اسماً ومقيداً وتعريفاً ، بالإضافة إلى رابطة حرة لقواعد المحتوى في شكل حقل يسمى قيم البيانات Data values يعرف بأنـه يتضمن توصيات باستخدام اللغات المقيدة أو القوائم المعيارية، كما تشتمل هذه المجموعة على روابط إلـى خطـة دبلن كور وغيرها من خطط "ما وراء البيانات" الأخرى. هكذا على النحو التالي :

RECORD TYPE
Qualifier: None
Definition: Identifies the record as being either a WORK record, for the physical or created object, or an IMAGE record, for the visual surrogates of such objects
Data Values (controlled) : work, image
VRA Core٢,٠: None
CDWA: None
Dublin Core: Type

ومن المثير للاهتمام أنـه لا يوجد معيار أو مواصفة قياسية حتى الآن لخطـط "مـا وراء البيانات" تحكم مـا ينبغـي أن تشتمل عليه تلك الخطط من

عناصر بيانات وتبين الكيفية التي ينبغي أن تمثل بها، وما هنالك سوى معيار الأيــزو لعناصــر البــيانات (ISO/IEC ١١١٧٩ specification and standardization of data elements) الـذي يقـرر عناصـر المعلـومات التـي يجب أن تتضمنها الملفات وقواعد البيانات(١). ويهدف من وراء هذا المعيار فـي الأسـاس إلـى ضـبط عناصـر البـيانات ليسهل فهمهـا وتـداولها . ويـضع الجـزء الأول مـن هـذا المعيار المكون من ستة أجزاء، إطاراً لفهم عناصـر البيانات ، وكيفية استخدام الأجزاء الأخـرى مـن المعيار، أمـا الجزء الثانـي فيعنـى باستخدام خطط التصنيف، ويحدد الجزء الثالث الخصائص الإجبارية والاختـيارية لوصـف عناصـر البـيانات، أمـا الجـزء الـرابع فيعطـي قـواعد وإرشـادات لكتابة تعريفات لعناصـر البيانات، ويركـز الجزء الخـامس علـى كيفية تحديد الأسماء والمعرفات ، بينما يخصص الجزء السادس لمسجلات registries عناصـر البيانات.

ونظـراً لأن عناصـر "مـا وراء البيانات" التـي تصف مصادر المعلومات تعد مجموعة فرعية من عـالم عناصـر البيانات ، فـإن هذا المعيار (ISO ١١١٧٩) ينبغي أن يطبق علـى الأقل علـى جزء من خطـة "ما وراء البيانات" ، وتحديداً الجزء الـذي يعرف عناصـر البيانات ، ويبدو أن مطوري خطط "مـا وراء البـيانات" المـتداولة فـي المكتـبات ومؤسـسات التـراث الفكـري لا

---

(١) These standards are available in print from the International Organization for Standardization (http://www.iso.ch). The best source for web-accessible versions and additional information about this family of standards is the European Commission's Diffuse Project (http://www.diffuse.org/). See Diffuse's summary page for ISO/IEC ١١١٧٩ at http://www.diffuse.org/meta.html #ISO ١١١٧٩.

يعون حقيقة الأمر جيداً، حيث يلاحظ أن جل الاهتمام بذلك المعيار في تلك المجتمعات كان من نصيب أولئك المهتمين بتطوير مسجلات "ما وراء البيانات" metadata registries .

مستويات الوصف :

تعمدت أن تكون الأجزاء السابقة غامضة أو غير مطردة بخصوص الشيء الذي تصفه "ما وراء البيانات" ؛ وذلك باستخدامي لمصطلحات مثل "كيان object "، و "شيء thing"، و "وحدة item " ، و"مصدر resource ". لأنه في واقع الأمر يمكن استخدام "ما وراء البيانات" لوصف عدة أنواع أو مستويات مختلفة من الكيانات تقع بين المفاهيم المعنوية المجردة إلى الموجودات المادية ، مع أن أحد الجوانب الأساسية لتعريف خطة "ما وراء البيانات" أو تعريف عنصر ما وراء البيانات هو تحديد نوع الكيان (أو الكيانات) الذي يمكن أن تطبق عليه الخطة أو عنصر البيانات.

وهناك نموذج معروف ومفيد لأنواع الكيانات التي يمكن وصفها ببليوجرافياً نجده في ما يعرف بالمتطلبات الوظيفية للتسجيلات الببليوجرافية للاتحاد الدولي لجمعيات المكتبات ومؤسساتها IFLA ، وهذا النموذج المعروف اختصاراً إف آر بي آر FRBR (والتي تنطق فر- بر fur-bur)[١]

---

(١) IFLA Study Group on the Functional Requirements for Bibliographic Records, Functional Requirements for Bibliographic Records: Final Report (Munich: K. G. Saur, ١٩٩٨), available at http://www.ifla.org/VII/sI٣/frbr/frbr.htm. Accessed I June ٢٠٠٢.

يحدد أربعة مستويات للكيان المادي، هي: العمل work ، وأسلوب التعبير expression، ووسيط العرض manifestation ، والوحدة item.

والعمل مفهوم معنوي مجرد يعرف بأنه "نتاج فكري أو فني متميز"؛ فمسرحية " عطيل Othello " لشكسبير مثلاً تعتبر عملاً ، وكذا أنشودة "المسيح Messiah " لهندل. والأعمال تعرف بأساليب التعبير عنها ، أو من خلال العروض المخصصة لها ، مثال ذلك إصدار طبعة خاصة لمسرحية " عطيل " ، أو لحن موسيقي خاص لأنشودة "المسيح" ، اللتين سبقت الإشارة إليهما، ويمكن أن يكون هناك - كما هو واقع الحال غالباً - العديد من أساليب التعبير للعمل الواحد ، مثال ذلك طبعاته المختلفة وترجماته وملخصاته والتوزيعات الموسيقية arrangements المرتبطة به ، مع ملاحظة أن أداء الأعمال الموسيقية والمسرحية يعد أحد أساليب التعبير مثله في ذلك مثل النوت الموسيقية والنصوص المرتبطة بها ، لكن التحوير الذي ينتج عنه أعمال فكرية أو فنية جديدة مميزة يعد عملاً جديداً ، وكذلك الحال عند تغيير العمل من شكل فني إلى شكل آخر؛ ومن هنا تعد أوبرا فيردي لمسرحية عطيل Othello عملاً قائماً بذاته ، له أساليب التعبير الخاصة به التي جاءت في شكل الألحان وكتابة الأوبرا والأداء.

أما وسائط العرض Manifestations فتعرف بأنها "التجسيد المادي لأسلوب التعبير الخاص بالعمل" ، أو جميع نسخ أساليب التعبير المنتجة على الوسيط نفسه وفي الشكل المادي نفسه. فمثلاً يمكن تسجيل أداء أوبرا فيردي لمسرحية عطيل في شكل فيلم ، وعلى أسطوانة مدمجة من نوع دي في دي DVD ، وعلى شريط فيديو من نوع في إتش إس VHS ، وعلى

أسطوانة مدمجة CD ، وعلى غيرها من الأشكال المتنوعة من الشرائط الصوتية ، فكل واحد من هذه الأشكال يمثل وسيط عرض مستقل ، وبالمثل يمكن تمثيل عمل نصي على وسائط عرض مختلفة كأن يطبع في نسخة مطبوعة بالحروف الطباعية العادية ، أو بالحروف الطباعية الكبيرة ، أو بتصويره على إحدى وسائط المصغرات الميكروفيلمية.

أما المكون الأخير في نموذج فربر فهو " الوحدة item" التي تعرف بأنها "نموذج وحيد لوسيط العرض " ، فهي عبارة عن كيان مادي واحد أو مجموعة كيانات مادية (مثال ذلك العمل الفني المسجل على اسطوانتين مدمجتين)، وبشكل عام يمكن القول أن جميع وحدات وسيط العرض الواحد لا تختلف بعضها عن بعض، مع توقع بعض " الأحداث الخارجة عن إرادة المنتج" والتي قد ينتج عنها ظهور أشكال أخرى مختلفة، كما هو الحال عندما تقوم المكتبة بإعادة تجليد بعض النسخ من كتاب معين، ومن ثم يكون لديها نسختان مختلفتان من الكتاب نفسه .

ويمكن أن يكون نموذج المتطلبات الوظيفية للتسجيلة الببليوجرافية مفيداً في المساعدة على توضيح ما نحاول أن نفعله مع ما وراء البيانات. فمثلاً القارئ الذي يبحث عن محتوى معين، لاشك أنه يهتم أولاً وأخيراً بالعثور على النسخة المناسبة من مصدر المعلومات الذي يضم بين ثناياه هذا المحتوى ، ثم يبحث في المقام الثاني عما إذا كانت تلك النسخة متاحة في شكل مطبوع أم غير ذلك ، ولاشك أن الوصف على مستوى أسلوب التعبير expression-level description له دوره في تلبية هذه الحاجة، ولكن من ناحية

أخرى قد لا يكون تحت تصرف القائم بإنشاء "ما وراء البيانات" سوى نسخة معينة وحيدة فقط من المصدر ومن ثم فالأفضلية ستكون لاعتماد الوصف على هذه النسخة، وهذا ما يؤدي إلى إيجاد نظام وصف على مستوى وسيط العرض manifestation-level description. ونظراً لأن المستفيدين من ذوي الاهتمامات الإدارية administrative users ، مثل القائمين بأعمال الحفظ Preservationists ، مهتمون بإدارة كيانات مادية معينة، نجد أن خطط "ما وراء البيانات" تميل من أجل تحقيق هذا الهدف إلى التركيز على مكون الوحدة item .

إن معظم خطط "ما وراء البيانات" لديها عناصر تخص أكثر من كيان من الكيانات التي يركز عليها نموذج المتطلبات الوظيفية للتسجيلة الببليوجرافية، وما لم تكن هناك نماذج بيانات صارمة تبين طبيعة الخطط ، فمن المتوقع حدوث قدر كبير من اللبس والتعقيد؛ فتسجيلة مارك٢١ الببليوجرافية - مثلاً - حددت مواضع للمعلومات المتعلقة لكل من الأعمال وأساليب التعبيرات ووسائط العرض والوحدات ، أما القواعد الأنجلو أمريكية للفهرسة فتدعو - مع بعض الاستثناءات - إلى إنشاء تسجيلة ببليوجرافية مستقلة لكل شكل من الأشكال التي يصدر بها العمل وبما يتناسب مع أي من وسائط العرض بمفهوم إفلا IFLA ، وقد أدى هذا إلى عدد من الإشكالات العملية لكل من المفهرسين ومستخدمي الفهارس عندما تكون هناك وسائط عرض متعددة، حيث يجب أن تكرر البيانات الببليوجرافية الخاصة بالعمل في كل تسجيلة مما يضيف بذلك إلى الفهرسة أعباءً جديدة ، بينما يواجه مستخدمو الفهارس مشكلة تعدد التسجيلات التي

يمكنها استرجاعها ، ويصبحون بذلك مجبرين على تخمين ما يميز كل تسجيلة عن الأخرى . وقد تم تقديم عدد من الحلول لمواجهة هذه المشكلة ، منها عمل تركيبة لتسجيلة متعددة الأجزاء multitiered record structure يتم فيها فصل عناصر البيانات التي تصف العمل عن عناصر البيانات التي تصف أسلوب التعبير ووسيط العرض .

وبقدر ما تتضح قيمة نموذج المتطلبات الوظيفية للتسجيلة الببليوجرافية FRBR عندما نفكر في إشكالات من هذا القبيل ، إلا أن هذا النموذج في حد ذاته لا يغطي جميع أنواع الكيانات التي يجب أن تتعامل معها "ما وراء البيانات"، كما أن هذا النموذج المقصور على الكيانات الموصوفة من قبل التسجيلة الببليوجرافية التقليدية ، لا يغطي كيانات أكبر ، من حيث مجال تغطيتها كالمجموعة Collection مثلاً، مع أن الوصف على مستوى المجموعة Collection-level description لا تقتصر أهميته فقط على المجموعة الأرشيفية (التي تم تعريفها من قبل منشئها provenance والمسؤولين عن صيانتها curators) ، وإنما على تجميعات aggregations المواد المتاحة عن طريق الشبكات ، والتي يمكن أن تشكل مجموعات اعتماداً على بعض الخصائص المحددة مثل تلك المجموعة التي يمكن عمل روابط لها من موقع واحد من مواقع الوب ، أو يمكن الوصول إليها عن طريق نظام استرجاع واحد ، أو التي تم وصفها عن طريق مجموعة وحيدة من ما وراء البيانات على مستوى الوحدة item-level metadata . وهناك مفهوم آخر له فائدة محتملة خاصة بالنسبة لأغراض إدارة الحقوق ،

وهو مفهوم " العمل الأكبر أو الأعلى superwork" الذي يقصد به جميع الأعمال التي انحدرت عن أصل مشترك ؛ فمسرحية عطيل لشكسبير وأوبرا عطيل لفيردي تنتميان لعائلة العمل الأعلى وهو " عطيل".

وتمثل المصادر الإلكترونية كذلك تحدياً لنموذج المتطلبات الوظيفية للتسجيلة الببليوجرافية، ذلك أن وسيط العرض يعرف بأنه جميع النسخ الصادرة من الوسيط نفسه وفي الصيغة المادية نفسها، ولكن ما يمثل ماهية الوسيط ترك مجهولاً، فهل يعتبر التخزين المباشر على القرص online disk وسيطاً؟ ، أم هل تعامل الأشكال المختلفة للتخزين على الأقراص - مثل الأقراص التي تعمل وفق معيار واجهة الحاسبات الآلية الصغيرة SCSI ، ومعيار البنية الفنية للتخزين المتسلسل SSA (١) - على أنها وسائط مختلفة؟ وهل يصبح القرص نفسه أو الشريط المادي ذاته وسيطاً مختلفاً عندما يتاح في إحدى المرات على الخط المباشر ، وفي مرة أخرى خارج الخط المباشر ؟ بداهة نقول إن هناك حداً معيناً من التمييز من المحتمل ألا يعرف بعده منشئو "ما وراء البيانات" شيئاً عن الاختلافات بين الوسائط ، ومن المحتمل كذلك إلا يهتم المستفيدون بعده كذلك بشيء عن تلك الاختلافات .

ومع ذلك ، فإن المحك هنا هو أن خطط "ما وراء البيانات" الجيدة يجب أن يكون لها نموذج واضح عن أنواع الكيانات التي تعنى بوصفها ، وعن طبيعة العلاقات المحتملة بينها. ومن المهم أيضاً أن ندرك (كما يفعل نموذج فربر FRBR) أن الكيانات غير الببليوجرافية لها علاقة بهذا الموضوع

(١) واجهة عالية السرعة طورتها شركة أي بي إم IBM [المترجم].

كذلك، فهناك أنواع أخرى من الكيانات مثل الوكلاء agents (الأشخاص والهيئات الاعتبارية) والأحداث والأماكن وحتى المعاملات. وغالباً ما تضم خطط "ما وراء البيانات" التي تركز على وصف المصادر، عناصر تخص كل نوع من أنواع تلك الكيانات؛ فتسجيلة الفهرسة التقليدية مثلاً تشتمل على بعض المعلومات عن الوكلاء (كما هو الحال بالنسبة لمؤلف الكتاب كاسمه وتاريخ ميلاده وتاريخ ووفاته)، وبعض المعلومات عن الأحداث (مثل تاريخ النشر ومكانه). ومع ذلك فإن الشخص الذي هو المؤلف يعتبر كياناً مستقلاً بذاته، ومن الممكن عمل تسجيلة "ما وراء بيانات" تصف ذلك الشخص بتفاصيل أكثر مما تسمح به تسجيلة الفهرسة، ومن ثم فإن عملية الربط بين تسجيلات "ما وراء البيانات" لأنواع مختلفة من الكيانات، ومع ما يثيره ذلك من إشكالات، لها أهميتها الكبرى التي لا تقل عن أهمية عملية ربط تسجيلات "ما وراء البيانات" بالمصادر المتصلة بها.

READINGS :

- Hodge, Gail. Metadata Made Simpler. Bethesda, Md.: NISO Press, ٢٠٠١. Available at http://www.niso.org/news/Metadata_simpler.pdf .

  A good general introduction to metadata schemes and issues, published in hard copy by NISO Press and available on the Web from the NISO website.

- IFLA Study Group on the Functional Requirements for Bibliographic Records. Functional Requirements for Bibliographic Records: Final Report (Munich: K. G. Saur, ١٩٩٨). Available at http://www.ifla.org/VII/sI٣/frbr/frbr.htm.

- Jones, Wayne, et al., eds. Cataloging the Web: Metadata, AACR, and MARC. ALCTS Papers on Library Collections and Technical Services no. ١٠ (Lanham, Md., and London: Scarecrow, ٢٠٠٢).

Short papers from the Pre conference on Metadata for Web Resources, ٦-٧ July ٢٠٠٠, at the ALA Annual Conference in Chicago.

- Tillett, Barbara B. "Bibliographic Relationships." In Relationships in the Organization of Knowledge, ed. Carole A. Bean and Rebecca Green (Dordrecht: Kluwer Academic Publishers, ٢٠٠١), ٩-٣٥.

More about FRBR.

There are several excellent clearinghouses of metadata schemes and related documents. This is by no means an inclusive list:

- The Canadian Heritage Information Network, CHIN, maintains a portal to metadata standards in several areas, with some emphasis on museum information

  (http://www.chin.gc.ca/Engiish/Standards/metadata-description.html).

- The Diffuse Project, funded by the European Commission's Information Society Technologies Program (http://www.diffuse.org/). This site maintains pointers to a large number of metadata-related standards and specifications in a wide range of subject areas.

- IFLA. maintains a listing of metadata schemes and related documents of interest to the library sector (http://www.ifla.org/II/metadata.htm).

MICI Metadata International Clearinghouse (Interactive), a site maintained by the Association of American Publishers (http://www.metadatainformation.org/) It focuses primarily on metadata of interest to publishers.

الفصل الثاني

تركيبة ما وراء البيانات : إنشاؤها واختزانها

تركيبة التبادل :

نبدأ مناقشتنا حول إدارة ما وراء البيانات بإلقاء نظرة على بعض الأشكال أو الصيغ المستخدمة بوجه عام لتمثيل ما وراء البيانات في شكل مقروء آلياً ؛ حيث يتم في بعض الحالات اختزان ما وراء البيانات ومعالجتها على نظم محلية بهذه الصيغ ، كما يتم في حالات أخرى أكثر تداولاً اختزانها في نظم قواعد بيانات محلية ، لكن تبادلها مع النظم الأخرى يتم عن طريق تلك الصيغ باعتبارها تركيبات نقل transport syntaxes ، وفي هذه الحالة سيحتاج النظام المحلي إلى تلقي ما وراء البيانات وإرسالها عن طريق صيغة واحدة أو أكثر من هذه الصيغ.

مارك **MARC :**

تعد صيغة مارك بالطبع من أكثر التركيبات تداولاً في مجال المكتبات ، وصيغة مارك MARC التي هي اختصار لعبارة Machine Readable Cataloging وتعني الفهرسة المقروءة آلياً ، طورتها مكتبة الكونجرس في منتصف الستينات من القرن العشرين الميلادي بهدف إنتاج بطاقات الفهرسة عن طريق الحاسب الآلي لتتمكن من توزيعها من خلال "خدمة توزيع الفهرسة"، ويعد استخدام صيغة مارك نقطة تحول جوهرية في تنظيم المكتبات من الناحية التقنية والتنظيمية ، كما أنها مهدت الطريق لظهور المرافق الببليوجرافية (مثل مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC ، وشبكة معلومات المكتبات البحثية RLIN ، وشبكة مكتبات

واشنطن WLN) ، وتبني مشروعات الفهرسة التعاونية الموسعة ، وهي التي ساعدت كذلك على تطوير النظم المتكاملة للمكتبات ، كما فتحت الباب أمام السوق التنافسية للنظم جاهزة التشغيل لإدارة المكتبات Turnkey Library Management Systems ، وتبعاً لذلك أدت الفهرسة التعاونية ومثيلاتها من نظم المكتبات بدورها إلى تعزيز الكيان الفكري للمكتبات من حيث الاتصالات والتعاون واحترام المعايير .

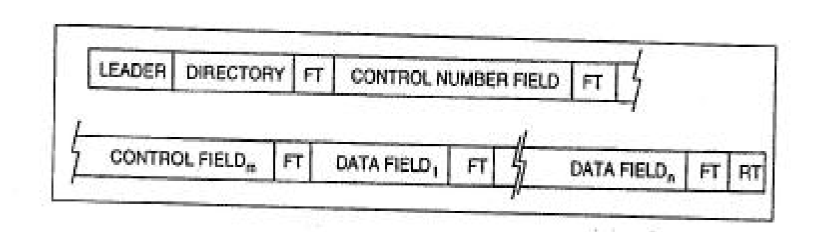
ومع ما أحدثته صيغة مارك من آثار وما تحظى به من نفوذ، ينبغي أن نعي أنها ليست خطة ما وراء البيانات ، ولكنها أحد المكونات المعقدة للخطة المتعددة الأوجه التي تشكل فهرسة المكتبات التقليدية ، ويأتي وصفنا لها بالمكون المعقد complex component ، من منطلق أن هذا المصطلح عندما يستخدم يقصد به شيئان على الأقل وبشكل عام ، أولهما البنية الخاصة للتسجيلات المقروءة آلياً المعرفة وفقاً للمعيار Z٣٩,٢ المعتمد من قبل كل من المعهد الوطني الأمريكي للمعايير والمنظمة الوطنية لمعايير المعلومات ANSI/NISO standard Z٣٩,٢ ، أما الآخر فمجموعة قواعد التكويد الموثقة في صيغة مارك ٢١ للبيانات الببليوجرافية MARC٢١ Format for Bibliographic Data ، وفي بعض المطبوعات الأخرى لمكتبة الكونجرس. (يلاحظ أن تسمية مارك الأمريكي USMARC عدلت إلى مارك ٢١ MARC٢١ في عام ١٩٩٨م كجزء من محاولة التقريب بين صيغ مارك الأمريكية والكندية والبريطانية).

ويحدد المعيار Z٣٩,٢ أن هذه الصيغة من صيغ تراسل البيانات تتكون من ثلاثة أجزاء ، الجزء الأول منها هو فاتح التسجيلة leader ، ويتكون من أربع وعشرين تمثيلة ، ويقسم إلى تسعة عناصر بيانات ، كل عنصر منها له معنى في شكل كود أو عدد. أما الجزء الثاني فهو دليل التسجيلة directory ، ويحتوي على شرائح معيارية entry طول كل منها اثنتا عشرة تمثيلة(١) ، وتمثل كل شريحة منها حقلاً من حقول البيانات بحيث توضح كل شريحة تاج الحقل، وطوله، وموضع بدايته. أما الجزء الثالث فيخصص لعدد متغير من الحقول، مع ملاحظة أن أياً من هذه الحقول إما أن يكون حقل ضبط controlled field (وتعرف بالحقول الثابتة) ، أو حقل بيانات data field(وتعرف بالحقول المتغيرة)، على أن يخصص لكل حقل ثابت عدد محدد مسبقًا من البيتات bytes مقسمة- كالفاتح - إلى عناصر بيانات ذات معاني خاصة بذلك الحقل. أما الحقل المتغير فيبدأ برمزين حجم كل منهما بايت واحد two one-byte flags وتعرف بالمؤشرات indicators ، يتبعها مضمون نصي مقسم إلى حقول فرعية تنتهي بحرف ختامي يعرف بفاصل الحقل terminator ، أما الحقول الفرعية فيرمز لها بحرف خاص يعرف بفاصل الحقل الفرعي delimiter (وهو حرف غير طباعي بنظام آسكي ويتمثل إما في شرطة رأسية أو علامة الدولار) يليه كود أو رمز من حرف واحد

---

(١) تقسم كل شريحة إلى ثلاثة عناصر ، أولها التاج ويخصص له ثلاث تمثيلات ، ثم طول الحقل ، ويخص له أربع تمثيلات ، وموقع تمثيلة بداية الحقل ويخصص له خمس تمثيلات [المترجم].

يشير إلى نوع الحقل الفرعي(١) ، وتُعرف كل من الحقول الثابتة والمتغيرة في الدليل بتاج مكون من ثلاث تمثيلات. يوضح الشكل رقم (١/٢) البنية الأساسية لصيغة مارك المكونة من : الفاتح ، والدليل ، وحقول الضبط ، وحقول البيانات.



الشكل رقم (١/٢) بنية تسجيلة مارك

المصدر: ANSI/NISO Z٣٩,٢ – ١٩٩٤(٢٠٠١) Information Interchange Format.

ملاحظة : تشير FT إلى فاصل الحقل ، و RT إلى فاصل التسجيلة

ومن الملاحظ أن اسم الحقل في بنية هذه التسجيلة لا يقع مجاوراً لبيانات الحقل، وإنما يوجد في دليل التسجيلة، ولكن عندما تعرض بيانات مارك على الشاشة أو عندما تطبع ، فإن تاج الحقل يظهر عادة كوسيمة أو علامة قبل بداية

(٢) تعرف بتقنينة الحقل الفرعي ، وهي عبارة عن رمز يستخدم لتحديد ذاتية الحقل الفرعي ؛ لأن الحقول المتغيرة قد تتكون من عنصر معلومات واحد أو أكثر من عنصر كما هو الحال مع حقل النشر الذي يتكون أساساً من مكان النشر واسم الناشر وتاريخ النشر ، وعادة تسبق تقنينة الحقل الفرعي كل عنصر معلومات في الحقل ، وتحد هوية هذا العنصر ، مع ملاحظة أن هذه الرموز لا تظهر في الحقول الفرعية عند عرض التسجيلة على الشاشة أو في الفهرس ، وإنما تستبدل بعلامة الترقيم المحددة في القواعد [المترجم].

بيانات الحقل، مع ملاحظة ورود المسافات قبل الحقول الفرعية وبعدها لتسهيل القراءة ، هكذا :

٢٤٥ ١٤$a The sound and the fury / $c by William Faulkner.

ومع أن معيار Z٣٩,٢ يحرص على توضيح معالم هذه البنية المعقدة إلى حد كبير، فإنه لا يذكر شيئًا عن المقصود بتيجان الحقول والمؤشرات والحقول الفرعية الممثلة في أي تسجيلة ببليوجرافية، وقد تكفلت بتعريف هذه العناصر الثلاثة المطبوعات الصادرة عن مكتبة الكونجرس ، وخاصة صيغة مارك ٢١ للبيانات الببليوجرافية MARC٢١ format for bibliographic Data ، والصيغة المختصرة منها التي تسمى " MARC٢١ Concise format for bibliographic Data " والمتاحة على الإنترنت عبر الموقع التالي: (http://www.loc.gov/marc/bibliographic/ecbdhome.html) ، كذلك يقرر هذا المعيار أن للحقل المتغير تاجاً يتكون من ثلاث تمثيلات[١] ، وتمثيلتين للمؤشرات ، وكل متغير من البيانات مقسم إلى حقول فرعية. كما تنص أدلة توثيق مارك٢١ على أن تاج حقل المدخل الرئيسي لأسماء الأشخاص هو التاج ١٠٠ ؛ وأن المؤشر الأول من هذا الحقل قد يأخذ القيم (٠) ، أو (١) ، أو (٣) على حسب طبيعة هذا الاسم ، أما المؤشر الثاني فغير محدد ، ومن ثم يجب أن يشغل بفراغ دائماً ، وتشير كذلك إلى أن الحقل الفرعي " a " يعبر عن اسم الشخص سواء في ترتيبه الطبيعي أو مبتدأ بلقب الأسرة أو

---

(١) قد تكون على هيئة حروف كبيرة (فقط) أو أرقام ، والتاج الوحيد المكون من أحرف هو حقل الفاتح leader وكان يشار إليه تبادليًا بالرمز (LDR) أو بالرمز (٠٠٠) [المترجم].

اسم الشهرة، ويوضح الشكل رقم (٢/٢) صفحة من صيغة مارك٢١ المختصرة للبيانات الببليوجرافية ، يظهر فيها استخدام الحقل رقم " ١٠٠ " .

ويشكل كل من المعيار Z٣٩,٢ ومطبوعات مارك٢١ تركيبة نقل transport syntax للمحتوى الذي يتم تسجيله في الفهارس التقليدية للمكتبات ، متضمنة ما يوضح الكيفية التي يجب أن تمثل بها البيانات لأغراض التبادل، وفي واقع الأمر تعد صيغة مارك٢١ توليفة مكونة من التركيبة والدلالات و قواعد المحتوى، حيث تُعرّف بمعاني أو دلالات كثير من عناصر البيانات التي لم تعرفها قواعد فهرسة المكتبات ، كما أنها تشتمل على القوائم الاستنادية التي تعتمد عليها في صياغة محتوى كثير من عناصر البيانات أو تحيل إلى هذه القوائم ، هذا فضلاً عن اشتمالها على قواعد المعالجة ، فالقيم التي تعطى للمؤشرات تستخدم غالباً لضبط عملية الفرز أو العرض أو غيرها من عمليات المعالجة المتصلة بمحتوى الحقل، مثال ذلك أن قيمة المؤشر قد ترمز إلى أن وسيمة معينة ينبغي أن تظهر قبل بيانات الحقل ، أو تشير إلى أن عدداً معيناً من الحروف يجب تجاهله عند الصف أو الترتيب.

ومن أوجه النقد التي توجه لصيغة مارك شدة تعقيدها، ذلك أن المعيار Z٣٩,٢ كان قد صمم في حقيقة الأمر ليعتمد عليه في تصميم نموذج قديم من نماذج معالجة البيانات، وكان معقداً دون مبرر، ولذلك فليس من المحتمل اليوم أن نصمم بنية تسجيلة بهذا التعقيد ، ومع ذلك فإن بنية المؤشر والحقل الفرعي لحقول البيانات لم تعد بهذا القدر من التعقيد إذا ما قورنت بطريقة تمثيل معلومات الحقول نفسها في صيغة لغة الترميز الموسعة XML

الحديثة، ولا شك أن تعقيد صيغة مارك يعود إلى الكم الهائل من المحتوى الذي تستطيع نقله ، وكذلك لتعقد قواعد المحتوى المستخدمة في هذا الصدد .

كما أن من أوجه النقد الأخرى الموجهة لصيغة مارك ، وهي على قدر كبير من الصحة، هي أنها صيغة منغلقة على نفسها a niche format ؛ حيث لم تستخدم إلا في داخل إطار المكتبات ، ولا تصحبها برمجية إعراب parsing ومعالجة متعددة الأغراض، ومع التسليم بصحة هذا النقد ، يجب ألا ننسى أنه لم يكن هناك قبل ظهور لغة الترميز الموسعة إلا النزر اليسير من تركيبات التسجيلات العامة واسعة الاستخدام التي تتعامل مع المعلومات النصية، ولا ننسى كذلك أنه عن طريق مارك تمكنت المكتبات لسنوات طوال من تبادل تسجيلات الفهرسة بعضها مع بعض من ناحية، ومع المرافق الببليوجرافية من ناحية ثانية، بل بقدر كبير من السهولة لم تستطع أن تحققه أية مجالات أخرى في تواصلها بعضها مع بعض .

MARC 21 Concise Bibliographic: Main Entry Fields (1XX)

[illegible]

- 100 - MAIN ENTRY--PERSONAL NAME (NR)
- 110 - MAIN ENTRY--CORPORATE NAME (NR)
- 111 - MAIN ENTRY--MEETING NAME (NR)
- 130 - MAIN ENTRY--UNIFORM TITLE (NR)

[illegible]
Items highlighted in red indicate changes made after the 2000 edition of the MARC 21 Concise Formats was published and are included in the 2001 printed edition of the MARC 21 Concise Formats

100 - MAIN ENTRY--PERSONAL NAME (NR)

[illegible]

Indicators

- First - Type of personal name entry element
  - 0 - Forename
    The name is a forename or is a name consisting of words, phrases, letters, etc., that is formatted in direct order.
  - 1 - Surname
    The name is a single or multiple surname formatted in inverted order or a single name without forenames that is known to be a surname.
  - 3 - Family name
    The name is a family, clan, dynasty, house, or other such group and may be formatted in direct or inverted order.
- Second - Undefined
  - # - Undefined

Subfield Codes

- $a - Personal name (NR)
  [illegible]
- $b - Numeration (NR)
  [illegible]
- $c - Titles and other words associated with a name (R)
- $d - Dates associated with a name (NR)

الشكل رقم (٢/٢)

صفحة من صيغة مارك ٢١ المختصرة للبيانات الببليوجرافية.

المصدر :

Library of Congress. Network Development and MARC standards office, MARC٢١ Concise Format for Bibliographic Data ٢٠٠١, concise edition. Available at http://www.loc.gov/marc/bibliographic/ecdbhome.html.

ويستهل مستند لغة ترميز النص الفائق بالتاج المعبر عن هذه اللغة وهو التاج <html> ، ويختتم بهذا التاج نفسه مسبوقاً بالشرطة المائلة هكذا: </html>، وبين هذين التاجين يقسم النص إلى قسمين رئيسين ، يعرف أولهما برأس النص ويرمز له بالتاج <head> ، ويعرف ثانيهما بمتن النص أو جسمه، ويرمز له بالتاج <body> ، على أن يظهر تاج العنوان <title> ، وتاج ما وراء البيانات <meta> الخاصة به في القسم الأول - أي رأس النص ، بينما يظهر المحتوى الفعلي لصفحة الوب في القسم الثاني - أي جسم النص أو متنه، ويتضح ذلك من خلال الشكل التالي لأحد المستندات المبسطة المكتوبة بهذه اللغة :

```
<HTML >
< HEAD>
<TITLE > Weather Report for Monday </TITLE >
<META NAME="title" CONTENT=" Weather Report for Monday" >
< HEAD>
```

```
<BODY >
< P>Warmer and slightly cloudy with a ٢٠٪ chance of afternoon
thunderstorms. </P >
 </BODY >
</ HTML>
```

وتجدر الإشارة إلى أن ما وراء البيانات يمكن تضمينها في مستندات لغة ترميز النص الفائق عن طريق استخدام التاج المسمى تاج ما وراء البيانات أو تاج الميتا<meta>. وهناك شكلان لكتابة هذا التاج الذي يلاحظ أنه تاج فردي يسجل محتوى البيانات بداخله وليس له تاج إغلاق هكذا :

```
< META HTTP-EQUIV= "text string١ " CONTENT="text string٢">
< META NAME= "text string١"  CONTENT="text string٢">
```

وواضح أن كلا الشكلين يقوم بتضمين ما وراء البيانات في المستند ، غير أن أولهما يعني أنه عند طلب المستند أو استدعائه، ينبغي على خادم الوب أن يستخدم ما يوجد مكان النص الأول ["text string١"] المحصور بين علامتي الاقتباس الأوليين كعنصر بيانات أساسي قابل للبحث ، ثم استخدام النص المحصور بين علامتي الاقتباس التاليتين "text string٢' كقيمة لهذا العنصر.

أما الشكل الثاني وهو الأكثر شيوعًا عند تسجيل ما وراء البيانات فيتم فيه تسجيل اسم عنصر ما وراء البيانات محل النص الأول المحصور بين علامتي الاقتباس الأوليين ["text string١"] ، وتسجيل القيمة المعطاة لهذا العنصر ، أي

الاسم محل النص "text string٢" المحصور بيت علامتي الاقتباس الأخيرتين . هكذا على النحو التالي:

< META NAME= "author" CONTENT="Smith, John">

مع ملاحظة إمكانية تضمين أي وسيمة label كقيمة لخاصية " الاسم"، إلا أن جدوى هذه الوسيمة ستعتمد على مدى قدرة نظم البحث على التعرف إليها، علماً بأن العديد من محركات البحث على الإنترنت تستطيع الآن أن تتعرف إلى بعض عناصر دبلن كور، كذلك يمكن برمجة محرك البحث المحلي ليتعرف إلى عناصر أي خطة من خطط ما وراء البيانات. ومع ذلك يوصى في هذا الصدد بأن يسبق اسم عنصر ما وراء البيانات ببادئة prefix تبين الخطة التي أخذ منها هذا العنصر ، ثم استخدام تاج الرابط <link> لربط هذه السابقة بأحد التعريفات المتاحة على الوب لهذه الخطة ، هكذا :

```
< META NAME= "DC. Creator" CONTENT="Smith, John">
< LINK REL= "schema. DC" HREF="http://
purl.org/DC/element/١,٠/">
```

وعليه يكون النموذج المكتمل لما وراء البيانات المضمنة في صيغة لغة ترميز النص الفائق للمثال السابق على النحو التالي :

```
<HTML >
<HEAD>
<TITLE> Weather Report for Monday </TITLE >
```

```
<META NAME="DC.Title" CONTENT="Weather Report for
Monday" >
<META NAME="DC.Creator" CONTENT="National Weather
Service" >
<META NAME="DC.Date" CONTENT="٢٠٠١-١٢-٠١" >
< LINK REL= "schema. DC" HREF="http://
purl.org/DC/element/١،٠/">
 </ HTML>
<BODY>
<P>Warner and slightly cloudy with a ٢٠٪ chance of afternoon
thunderstorms.</P>
</BODY>
</ HTML>
```

لغة الترميز المعيارية العامة **SGML** :

المستند في لغة ترميز النصوص الفائقة HTML هو في واقع الأمر تطبيق للغة الترميز المعيارية العامة SGML (أحد المعايير التي أقرتها هيئة المواصفات الدولية " أيزو " واسمه ISO ٨٨٧٩:١٩٨٦ Information processing- Text and Office systems- Standard Generalized Markup Language.). ولغة الترميز المعيارية العامة تعد بشكل رسمي لغة متعدية metalanguage أو لغة لوصف لغات أخرى، وهي تحدد مجموعة من قواعد تركيب نوعية لترميز الوثائق أو المستندات (كما هو الحال مع قاعدة كتابة

أسماء التيجان بين قوسين زاويين)، لكنها لا تحدد أية مجموعة تيجان معينة، وإنما تتيح لأي شخص حرية تحديد مجموعة تيجان خاصة به مع قواعد استخدامها، ويتم ذلك من خلال إنشاء معرف نوع المستند أو الوثيقة document type definition ، أو ما يسمى اختصاراً دي تي دي DTD ، ومن الملاحظ أن معرف نوع الوثيقة قد يسمى - على سبيل المثال - لغة ترميز النص الفائق HTML ، وينص فيه على أن التيجان المعتمدة allowable تشمل التيجان التالية على سبيل المثال : تاج العنوان <title> ، و تاج ما رواء البيانات <meta> ، وتاج الربط <link> ، و تاج رأس النص <head> ، و تاج الفقرة <p> ، وأن كلاً من تاج العنوان <title> ، و تاج ما وراء البيانات <meta> ، وتاج الرابط <link> ، ينبغي أن تسجل في الجزء الخاص برأس الصفحة <head> ، وليس في مكان آخر ، أما تاج الفقرة <p> فيجب أن يكتب فقط في الجزء الخاص بجسم النص <body>.

وكما اتضح لنا من خلال استعراضنا للغة ترميز النص الفائق فإن لغة الترميز المعيارية العامة تقوم بترميز بعض محتوى البيانات بين تاجي البداية والنهاية ، ويكود البعض الآخر من محتوى البيانات على أنه قيم " للخصائص attributes " التي تلي اسم تاج البداية ، والمثال على ذلك تاج ما وراء البيانات <meta> الموضح في المثال التالي ، الذي له الخاصية المسماة " الاسم name " التي تأتي قيمتها بعد علامة يساوي (=) ، هكذا :

```
< META NAME= "title" …>
```

وينبغي هنا أن تحدد طبيعة أي من هذه الخصائص ، بمعنى هل هي اختيارية أم إجبارية ، كما ينبغي أن تعطى قائمة بقيم الخصائص المعتمدة في هذا الصدد ، وكذلك تحدد هذه اللغة طبيعة البيانات التي يتضمنها كل عنصر من عناصر لغة الترميز المعيارية العامة ، بمعنى هل هذا العنصر يتضمن بيانات نصية مع خاصية واحدة أو أكثر من خاصية ، أم هل يتضمن بيانات نصية فقط، أو هل يتضمن خصائص فقط ، كذلك يمكن أن تحدد العناصر التي تتخذ من العناصر الأخرى محتوى لها، وفي هذه الحالة يسمى العنصر الخارجي أو المتضمِن لهذه العناصر " بالغلاف wrapper" ، بينما تسمى العناصر المتضَمنة nested بداخل حدود هذا العنصر بالعناصر الفرعية subelements"، مثال ذلك عنصر التاريخ <date> الذي يعد غلافاً عندما يشتمل على كل من : الشهر <month> ، واليوم <day> ، والسنة <year> كعناصر فرعية على النحو التالي :

```
<date>
<month>٠٢
<day>١٢
<year>٢٠٠٢
</date>
```

كذلك من الممكن ألا يتضمن العنصر أية بيانات ، سواء أكانت بيانات نصية أو عناصر فرعية ، وفي هذه الحالة يسمى هذا العنصر " بالعنصر الفارغ أو الخالي empty element " ، ومن أمثلة ذلك الخط الذي يوضع في سياق أحد النصوص line break ، فهذا الخط يمكن أن يشار إليه على أنه

عنصر خال ، حيث يتم إدراج التاج المعبر عنه (<lb>) في الموضع الذي يراد أن يوضع فيه هذا الخط ، هكذا:

<lb> </lb>

وبشكل عام يمكن القول إن لغة الترميز المعيارية العامة تعد من نواحٍ كثيرة تركيبة تكويد جيدة لما وراء البيانات encoding syntax ؛ فقد صممت لتتعامل مع البيانات النصية متغيرة الأطوال، حيث يمكن تعريف عدد لا حدود له من العناصر (أي التيجان والخصائص) بشرط أن تكون أسماؤها واصفة لمحتوياتها؛ فإذا كان الوصف الأرشيفي المرمز EAD - مثلاً - يستخدم التاج <titleproper> للتعبير عن عنصر العنوان نفسه title proper ، فإنه بالتأكيد أثبت في الذاكرة من الحقل المخصص لذلك في صيغة مارك ، وهو الحقل الفرعي a التابع للحقل رقم " ٢٤٥ "، كما أن هذه اللغة يمكن أن نعدها لغة طبقية أو هرمية hierarical بطبيعتها ، ويمكن أن تنفذ هذه الطبقية مما يجعلها وسيلة جيدة للتعبير عن أنواع العلاقات الطبقية الكائنة بين المجموعات من ناحية ، وبين كل من الأعمال وأساليب التعبير ووسائل العرض والوحدات من ناحية أخرى ، كما أن خاصية التوارث التي تكتسبها هذه اللغة تجيز للعناصر من مستوى طبقي أقل أن ترث المعلومات المرمزة في المستويات الأعلى منها ، وبالتالي لا تكون هناك حاجة لتكرار البيانات الوصفية، يضاف إلى ذلك أن قدرة عناصر هذه اللغة على احتواء عناصر أخرى تتيح إمكانية تحديد عناصر ما وراء البيانات بطرق مرنة للغاية، ففي الجزء التالي من وصف إحدى معينات البحث بصيغة الوصف الأرشيفي

المرمز EAD، يلاحظ أن الجزء الخاص بالعنوان (<titlestmt>) يحتوي على عنصري العنوان نفسه والمؤلف ، هكذا:

```
<titlestmt>
<titleproper>A Guide to the Paul A. M. Dirac
Collection</titleproper>
<author> Processed by Brt Altman and Charles Conaway</author>
</titlestmt>
```

ونظراً لأن معرف نوع الوثيقة في صيغة الوصف الأرشيفي المرمز EAD DTD يسمح بذلك ؛ فمن الممكن إذاً وضع محددات لبعض الأسماء الشخصية - مثلاً - حتى يمكن معالجتها ، وخصوصاً لأغراض التكشيف أو العرض على الشاشة ، هكذا :

```
<titlestmt>
<titleproper>A Guide to the <persname> Paul A. M. Dirac
</persname> Collection </titleproper>
<author> Processed by Brt Altman and Charles Conaway </author>
</titlestmt>
```

أما فيما يتصل بدلالات أو معاني semantics عناصر أية خطة من خطط من وراء البيانات فيمكن عرضها في صيغة لغة الترميز المعيارية العامة عن طريق إنشاء معرف مناسب من معرفات نوع الوثيقة DTD ، وهذا بالفعل ما قامت به مكتبة الكونجرس حيث أنشأت معرفاً لنوع الوثيقة خاصاً بمارك ٢١ في صيغة اللغة المعيارية الموحد للترميز ، مع المحافظة على جميع تيجان مارك٢١.

وباستخدام هذا المعرف يمكن للعنوان التالي المرمز في صيغة مارك٢١ هكذا:

٢٤٥ ٠٤ $a The health of diary cattle/ $c edited by Anthony H. Andrews.

أن يرمز بصيغة لغة الترميز المعيارية العامة على النحو التالي:

<mrcb٢٤٥ il=٠ i٢=٤><mrcb٢٤٥-a>The health of diary cattle/<mrcb٢٤٥-c>edited by Anthony H. Andrews.

وهنالك بالفعل عدد من خطط ما وراء البيانات وضعت لها معرفات من قبيل معرفات أنواع الوثائق بصيغة لغة الترميز المعيارية العامة " SGML DTD" ، وهذا يعني أنه بدلاً من استخدام الأسماء المستقلة التركيب syntax-independent names والتعريفات الخاصة بكل عنصر من عناصر ما وراء البيانات تُعرّف العناصر عن طريق تيجان وخصائص لغة الترميز المعيارية العامة . وعندما يتم تعريف خطة ما وراء البيانات عن طريق معرف نوع الوثيقة يأتي دور الأدوات المساندة في هذا الصدد؛ ذلك إننا لا نتوقع من كل من هو معني بوصف المصادر أن يكون مستوعباً لمعرف نوع الوثيقة. ويعد دليل أو مكتبة التيجان[١] tag library من أكثر الأدوات إفادة في هذا الصدد، وهذه المكتبة عبارة عن وثيقة تشتمل على قائمة مرتبة هجائياً بجميع أسماء عناصر لغة الترميز المعيارية العامة وخصائصها، مع

(١) لا تستخدم مع لغة الترميز المعيارية العامة فقط ، وإنما مع لغة الترميز الموسعة كذلك [المترجم].

تعريفاتها، مصحوبة بقواعد يمكن للبشر استيعابها لاستخدام هذه التيجان. ويبين الشكل رقم (٣/٢) أحد مداخل المكتبة التاجية للوصف الأرشيفي المرمز EAD في الإصدارة رقم ١,٠.

الشكل رقم (٣/٢)

مثال لمكتبة التيجان في صيغة لغة الترميز الموسعة.

المصدر :

the Encoded Archival Description: Tag Library, version ١,٠. Chicago: the society of American Archivests,١٩٨٨.

<div> Text Division

Description:

A generic element that designates a major section of text within <frontmatter>. Examples of these text divisions include a title page, preface, acknowledgements, or instructions for using a finding aid. Use the <head> element to identify the <div>'s purpose.

May contain:

address, blockquote, chronlist, div, head, list, note, p, table

May occur within:

div, frontmatter

Attributes:

ALTRENDER	#IMPLIED, CDATA
AUDIENCE	#IMPLIED, external, internal
ID	#IMPLIED, ID

Example:

```
<frontmatter>
<titlepage>
      [other possible elements and text... ]
</titlepage>
<div><head>Acknowledgements</head>
<p>The University of Ithaka Manuscript Library wishes to acknowledge the
generosity of the family of Edgar Holden for an endowment in support of
the processing and maintenance of this collection and the Clark Foundation
for grant funding in support of the encoding of this and other finding
aids using the Encoded Archival Description standard.</p></div>
      [other possible elements and text... ]
</frontmatter>
```

لغة الترميز القابلة للتوسع أو الموسعة **XML** :

مع ما تتميز به لغة الترميز المعيارية العامة SGML من جوانب قوة متعددة ، إلا أن ثمة نقطة ضعف بارزة تكتنفها ، وهي أنها لغة يصعب على البرامج معالجتها؛ ولذلك فلم يصمم إلا القليل نسبياً من البرامج التطبيقية التي تدعم إنشاء النصوص المكودة بهذه اللغة واختزانها وتهيئتها وعرضها ، بل إن ما يتوافر من تلك البرامج يتسم بالتعقيد . ويمكن النظر إلى لغة الترميز الموسعة على أنها مجموعة فرعية من لغة الترميز المعيارية العامة ، ولكن وفق قواعد أكثر تحديداً وخصائص و خيارات أقل حتى يسهل معالجتها؛ فإذا كان من الممكن في سياق اللغة المعيارية الموحد للترميز حذف تيجان النهاية في حالات معينة أو حصر قيم الخصائص بين علامتي الاقتباس أو عدم حصرها، ففي سياق لغة الترميز الموسعة و إذا كان لأحد العناصر تاج نهاية ، ينبغي أن يستخدم هذا التاج ، وينبغي كذلك أن تكتب قيمة الخاصية بين علامتي الاقتباس ، وبناءً على ذلك فإذا أردنا ترميز المثال السابق ولكن في صيغة لغة الترميز الموسعة فإن ذلك يتم بطريقة مختلفة هكذا على النحو التالي:

```
<mrcn٢٤٥ il="٠"><mrcb٢٤٥-a>the health of diary
cattle/</mrcb٢٤٥a>
<mrcb٢٤٥-c>edited by Anthony H. Andrews.</mrcb٢٤٥-
c></mrcb٢٤٥>
```

وهناك بعض الاختلافات الأخرى التي قصد من ورائها تطويع لغة الترميز الموسعة بشكل أكبر لتتلاءم وطبيعة الشبكة العنكبوتية العالمية ، ولعل من أكبر الأمثلة على ذلك أن تكون تيجان هذه اللغة مقيدة بحالة الحروف[1] (وعليه فإن التاج <META> يعد تاجًا مختلفاً إذا كتب هكذا <meta>) ، يضاف إلى ذلك أن هذه اللغة تشتمل على عدد من المحارف الأخرى التي لا يتضمنها نظام آسكي للرموز الموحدة ASCII UNICODE ، مما يجعل تكييف استخدامها يتم بشكل أفضل في البيئات الدولية غير الملتزمة باستخدام الأبجدية الرومانية.

إن تطوير هذه اللغة وما يتصل بها من مواصفات كان من ثمار إحدى المبادرات النشطة التي تبناها ائتلاف شبكة الوب (W٣C)[٢] . ومن أهم تلك المواصفات المتصلة بها والتي أقرها هذا الائتلاف ، والتي لا تزال قيد التطوير مواصفة لغة الربط في صيغة لغة الترميز الموسعة XML linking language المعروفة اختصاراً " XLink " ، ومواصفة لغة الاستفسارات في صيغة لغة الترميز الموسعة XML query language ، المعروفة اختصاراً " XQuery " ومواصفة إخراج أو عرض الصفحات a stylesheet specification (التي تعرف بلغة عرض الصفحات الموسعة Extensible Stylesheet

---

(١) تقصد بالرسم الإملائي الذي تكتب به الحروف ، بمعنى أن الحرف المكتوب بالشكل الكبير منه يختلف عندما يكتب بالشكل الصغير [المترجم].

(٢) Extensible Markup Language (XML) (home page of the W٣C XML Activity), available at http://www.w٣.org/XMU. Accessed ٣ June ٢٠٠٢.

Language، المعروفة اختصاراً "XSL ")، ولغة ترميز النص الفائق الموسعة XHTML ، التي تعد بمثابة إعادة صياغة للغة ترميز النص الفائق ليتوافق مع لغة الترميز الموسعة. كذلك من أهم الفعاليات ذات الصلة بهذه اللغة تطوير مخطط أو خطة لغة الترميز الموسعة XML schema التي ينظر إليها على أنها طريقة بديلة لمعرفات أنواع الوثائق ،حيث تدعم المزيد من القواعد البنائية والمزيد من إجراءات التحقق من المحتوى content validation أكثر مما كانت تقوم به معرفات أنواع الوثائق ، وفي هذا الصدد قامت مكتبة الكونجرس بتطوير خطة ترميز موسعة خاصة بمارك ٢١ ، تعرف بمارك إكس إم إل MARCXML ، واعتماداً على هذه الخطة يرمز حقل العنوان في المثال السابق على النحو التالي:

```
<datafield tag="٢٤٥" indl ="٠" ind٢="٤">
<subfield code="a"> the health of diary cattle/</subfield>
<subfield code="c">edited by Anthony H. Andrews.</subfield>
  </datafield>
```

لقد اكتسبت لغة الترميز الموسعة مكانة مهمة حتى أصبحت تعرف بلغة الوب، وباتت متصفحات الوب تتضمن ما يدعم الوثائق المكودة بهذه اللغة ، كما استحدثت العديد من البرمجيات التي تدعم جهود التكويد والتخزين والاسترجاع والعرض بهذه اللغة، كما تم تعيين بعض خطط ما وراء

البيانات المطورة حديثاً جداً كمبادرة توثيق البيانات[1] Data Documentation Initiative على أنها من قبيل معرفات أنواع الوثائق في صيغة لغة الترميز الموسعة (XML DTDs)، وفضلاً عن ذلك حولت بعض الخطط التي كانت قد طورت أساساً كمعرفات أنواع الوثائق بلغة الترميز المعيارية العامة (SGML DTDs) ، لتستخدم كمعرفات أنواع الوثائق بلغة الترميز الموسعة (XML DTDs). ومن المحتمل أن يتم تعريف خطط ما وراء البيانات المستقبلية باستخدام خطط لغة الترميز الموسعة بدلاً من معرفات أنواع الوثائق DTDs.

إطار الوصف المصدري أو وصف المصدر :

إطار وصف المصدر Resource Description Framework ، أو ما يسمى اختصاراً آر دي إف RDF ، هو محاولة لتمثيل ما وراء البيانات على الوب بالقدر الذي لا يكفل لها أن تكون مقروءة آلياً فحسب ، وإنما لتكون مفهومة آلياً كذلك. ويعد إطار وصف المصدر رسميًا نموذج بيانات لتمثيل كل من : المصادر، وسمات هذه المصادر ، وقيم تلك السمات. ويمكن نظرياً التعبير عن هذا النموذج بأكثر من تركيبة واحدة. ومع ذلك فإننا عندما نتأمل إطار وصف المصدر، نتأمل عموماً تمثيله في لغة الترميز الموسعة ، ويؤكد ذلك أن الجزء الأكبر من مواصفة هذه الإطار يعنى بقواعد التعبير عن (أو

(١) مبادرة توثيق البيانات DDI ، منظمة منتجي بيانات العلوم الاجتماعية والأرشيفيين في هذا المجال ، ويشير المصطلح كذلك إلى مواصفة ما وراء البيانات الذي طورته هذه المبادرة لوصف مجموعات العلوم الاجتماعية [المترجم].

"تتابع serializing ") نموذج البيانات الخاص به في صيغة لغة الترميز الموسعة(١).

ومن الأفكار الأساسية المرتبطة بإطار وصف المصدر فكرة الأسماء الفريدة أو المميزة namespaces التي ينظر إليها وكأنها أشبه بالإصدارات المتاحة على الوب من خطة ما وراء بيانات. وتقوم الفكرة على أن أي عنصر ما وراء بيانات يرد في سياق أي وصف يعتمد على إطار وصف المصدر RDF Description ينبغي أن تسبقه وسيمة label تربطه باسم مميز معين، وهذا الأسلوب كفيل بأن يحقق هدفين ، يتمثل أولهما في أن اسم العنصر سترافقه أو ترتبط معه إحدى سبل الحصول على تعريفه، ويتمثل الآخر في إمكانية استخدام عدة عناصر من أكثر من خطة من خطط ما وراء بيانات دون أن يحدث ذلك أي نوع من التضارب عند وصف مصدر واحد.

وبناءً على ذلك فإن وصف أحد تقارير الأحوال الجوية اعتماداً على خطة دبلن كور ، الذي عرض في فقرة سابقة في صيغة لغة ترميز النص الفائق، يمكن تمثيله في صيغة إطار وصف المصدر على النحو التالي:

```
<?xml version="١,٠"?>
<!DOCTYPE rdf:RDF SYSTEM
http://dublincore.org/documents/٢٠٠١/٠٤/١١/dcmes-xml/dcmes-xml-dtd.dtd">
```

(١) Resource Description Framework (RDF) Model and Syntax Specification, W٣C Recommendation ٢٢ February ١٩٩٩, available at http://www.w٣.org/TR/REC-rdf-syntax/ Accessed ٣ June ٢٠٠٢.

```
<rdf:RDF xmlnsdrf=http://www.w٣.org/١٩٩٩/٠٢/٢٢-rdf-syntax-ns#
xmlns:dc="http//ourl.org/dc/elements/١,١">
<rdf:Discription about="http://[URL of weather report page]">
<dc:title> Weather Report for Monday<dc:title>
<dc:creator> National Weather Service</dc:creator>
<dc:date>٢٠٠١-١٢-٠١</dcdate>
</rdf:Description>
</rdf:RDF>
```

ومن الملاحظ أن تسجيلة إطار وصف المصدر تبدأ بإعلان اسمين مميزين هما : مواصفة إطار وصف المصدر نفسه ، ودبلن كور ، مع ربط الأول بالوسيمة "rdf"، وربط الثاني بالوسيمة "dc" ، وفي هذا إشارة إلى أن اسم أي عنصر تسبقه الوسيمة "rdf " يكون معرفاً في المستند المتاح على العنوان التالي : http://www.w٣.org/١٩٩٩/٠٢/٢٢-rdf-syntax-ns# ، وبالمثل فإن اسم أي عنصر يسبق بالوسيمة "dc " سيكون معرفاً في المستند المتاح على العنوان التالي : http://purl.org.dc/elements.١,١/.

ثم يلي ذلك عنصر الوصف <description> الخاص بإطار وصف المصدر والذي يجمع البيانات المتصلة بمصدر واحد ، وتبدو فيه خاصية الحيثية أو " عن about " التي تعرف المصدر بتوضيح محدده identifier ، وهو - في هذا المثال - المحدد العام لموقع المصدر URL [١]، ثم تأتي بعد

(١) تقصد المعرف التالي : "http://[URL of weather report page]" [المترجم].

ذلك ثلاثة عناصر بيانات من عناصر دبلن كور ، كل منها يسبق بالوسيمة "dc" - الدال بالطبع على خطة دبلن كور.

ومع أنه بالإمكان تضمين إطار وصف المصدر داخل المستند في صيغة لغة ترميز النص الفائق ، إلا أن المعالجة نفسها التي يقوم بها المتصفح الأساسي ستتجاهله ، وإن كانت هناك بعض المتصفحات التي تحاول عرض المحتوى. وما يشيع الآن هو عمل إطار وصف المصدر كمستند خارجي قائم بذاته، ومن ثم تقوم صفحة لغة ترميز النص الفائق بالإحالة إلى وصف إطار وصف المصدر الخاص بها عن طريق تاج الرابط <link> هكذا على النحو التالي:

```
<HTML>
<HEAD>
<TITLE>weather Report of Monday </TITLE>
<LINK REL="meta" HRE"[URL of RDF description]">
</HEAD>
<BODY>
<P>warner and slightly cloudy with a ٢٠٪ chance of afternoon
thunderstorms.</P>
</BODY>
</HTML>
```

ومع أن إطار وصف المصدر يضيف قيوداً بنائية إضافية على لغة الترميز الموسعة ، إلا أن هذه القيود ينظر إليها البعض على أنها من المزايا، في الوقت الذي يعدها آخرون غير ذلك - أي عيوب ، غير أن هذا يتوقف

على ما يطمح المرء إلى إنجازه. كذلك يمكن القول إن إطار وصف المصدر يعيبه كثرة الإطناب، ومع ذلك تفضل بعض تطبيقات ما وراء البيانات تحديد تركيبة النقل الخاصة بها في صيغة إطار وصف المصدر مستثمرة بذلك إمكانية استيعابه آلياًmachine-understandability ، هذا في الوقت الذي لا تزال تفضل فيه تطبيقات أخرى أن ترد هذه التركيبة في صيغة لغة الترميز الموسعة البسيطة.

إنشاء ما وراء البيانات واختزانها :

عندما يقوم البشر بإنشاء ما وراء بيانات فإنهم عادة ما يستعينون في ذلك ببعض البرمجيات، أما في عالم المكتبات - وحيث يعد إنشاء ما وراء البيانات وفقاً للقواعد الأنجلو أمريكية للفهرسة ومارك AACR٢/MARC أمراً مهنياً مدروساً - فهنالك نموذجان أساسيان لمعالجة هذه العملية ، النموذج الأول يتم من خلال وصف المصادر عبر نظم الفهرسة التعاونية التي يرعاها أحد المرافق الببليوجرافية، مثل مركز الحاسب الآلي للمكتبات على الخط المباشر المعروف بـ (أو سي إل سي OCLC) ، و شبكة معلومات المكتبات البحثية RLIN ، اللذين يتبنيان إنشاء تسجيلات مارك بناء على قواعد الفهرسة الأنجلو أمريكية في طبعتها الثانية MARC/ AACR٢ ، كذلك ما يقوم به أول هذين المركزين (أو سي إل سي OCLC) من دعمه مجموعة دلالات ما وراء البيانات التي يتم تمثيلها إما في صيغة مارك أو دبلن كور عبر نظام الفهرس التعاوني للمصادر على الإنترنت والمعروف بكورك CORC, cooperative online Resource Catalog (هذا الفهرس الآن جزء من نظام أو سي إل سي للفهرسة والمعروف ب OCLC Connexion).

وبحكم ما توافر لهذه النظم من قواعد بيانات مركزية ضخمة قامت على تجميعها على مدار فترات طويلة ، فمن الممكن البحث عن التسجيلات الخاصة بأحد المصادر محل الوصف (والتي غالباً ما يعثر عليها) ، ثم تقوم إحدى المكتبات التي تبحث عنها باستنساخها كما هي أو بشيء من التعديلات، ومن ثم فإن التسجيلات التي أنشئت بهذه الطريقة يمكن استيرادها من النظام المركزي ، كما يمكن تزويد النظام المحلي للمكتبة بها لعرضها في الفهرس المحلي الخاص بها ، ولا شك أن هذا النمط من الفهرسة والمعروف بالفهرسة التعاونية يوفر الوقت والخبرات والنفقات اللازمة لإنشاء التسجيلات، كما أنه يقلل كذلك من أخطاء إدخال البيانات.

أما في النموذج الثاني فيتم فيه إنشاء التسجيلات عبر النظام المتكامل للمكتبة (ILS) ، ثم يتم بعد ذلك المشاركة بها في أحد المرافق الببليوجرافية ، أو قد لا يتم ذلك ، وعليه يحرص معظم وكلاء النظم الآلية المتكاملة على تضمينها وظائف دعم إنشاء تسجيلات مارك داخلياً وإمكانات تعديلها وتحريرها، بل إن بعض الموردين بدأوا بالفعل إضافة وظائف دعم خطط ما وراء البيانات، ذلك أن واجهات المستفيدين في كل من النظم المتكاملة الداخلية والمرافق الببليوجرافية معقدة نسبياً وخاصة ما يتصل بعمليات إدخال بيانات مارك والتحقق من صحتها أو على الأقل تحديد محتوى تلك البيانات، حيث تستطيع البرامج - مثلاً - التأكد من أن حقلاً ما لا يحتوي إلا على قيم مؤشرات صحيحة والتأكد من صحة ترميز الحقل الفرعي، ومع أن البرامج لا تستطيع التأكد من صحة النص المدخل ومدى الالتزام بمجموعة القواعد المناسبة في ذلك ، تتميز العديد من النظم المتكاملة للمكتبات بقدرات -

وإن كانت معقدة إلى حد ما - على التدقيق الإملائي والتثبت من مدى استنادية رؤوس الموضوعات ، وخاصة للرؤوس الخاصة بالأسماء والموضوعات، اعتماداً على مضاهاتها بالملفات الاستنادية المناسبة .

ولأنه يندر أن تختزن ما وراء بيانات مارك وفقاً لصيغة Z٣٩,٢ ، تقوم معظم النظم بتحليل عناصر البيانات إلى جداول قواعد بيانات علاقية، ثم بناء كلمات أساسية مستقلة وكشافات للمصطلحات string indexes من البيانات الأساسية لتحسين كفاءة استرجاع البيانات، ومع ذلك فإن نظم المكتبات تعي تماماً بنية مارك، كما أن إجراءات المعالجة الخاصة الواعية بمدلولات قيم المؤشرات والحقول الفرعية تعد شيئاً مألوفاً بالنسبة لهذه النظم .

ومع أن النظم المكتبية تدعم إنشاء وصيانة تسجيلات مارك من عقود مضت، إلا أن الحاجة لإنشاء صيغ أخرى من ما وراء البيانات من الأمور الجديدة نسبيًا على تلك النظم ، كما يحيط بالنظم التي تدعم هذه الصيغ قلة الإتقان والتنوع الشديد، ربما لأن العديد من تلك النظم طورت على مستوى محلي وليس عن طريق الشراء من الوكلاء التجاريين ، ومع أنه يمكن استخدام أدوات التحرير العامة لكل من لغة الترميز المعيارية العامة ولغة الترميز الموسعة في هذا الصدد أيضاً، أو استخدام الإصدارات الجديدة من برامج معالجة النصوص متعددة الأغراض التي تدعم مستندات لغة الترميز الموسعة ، فإن الأفضلية عادة ما تكون من نصيب استخدام برامج التركيب أو التجميع الخاصة special authoring programs ، لاسيما وأن الجيل الحالي من أدوات التركيب يميل إلى التعامل مع لغة الترميز المعيارية العامة ولغة

الترميز الموسعة ، ويوفر مجموعة وظائف كاستخدام معرفات أنواع الوثائق أو المستندات لإنشاء النماذج الجاهزة template لإدخال البيانات ، وإمكانات التثبت من أن البيانات المدخلة تتوافق مع صيغة لغة الترميز الموسعة توافقاً كاملاً ، والتثبت كذلك من أن البيانات المدخلة صحيحة (أي تتوافق مع معرف نوع الوثيقة) ، هذا فضلاً عن عرض الترميز الداخلي nested markup بشكل مناسب ، واستخدام نماذج عرض أو إخراج الصفحات stylesheet لضبط طريقة عرض البيانات.

وهناك أسلوب عام آخر لتركيبات أكثر تبسيطاً لما وراء البيانات يتمثل في استخدام نموذج وب web form كنموذج جاهز لإدخال البيانات ، حيث يتم التحقق من صحة البيانات ثم تحويلها بعد ذلك إلى صيغة تخزين ، كما هو الحال عندما يتم ضبط الصفوف في جداول قواعد البيانات العلائقية مثلاً. وقد تكفلت بعض مبادرات ما وراء البيانات بتطوير أدوات إدخال بيانات، أو دعت لاستخدام أدوات طورتها أطراف أخرى مستقلة، إلا أن كثيراً من هذه الأدوات يعد مناسباً في مرحلة الإنشاء الأولية للبيانات أكثر من ملاءمته لإجراءات الصيانة المستمرة للبيانات مع مرور الوقت.

وبمجرد إنشاء ما وراء البيانات ، سواء في صيغة لغة الترميز المعيارية العامة أو لغة الترميز الموسعة ، ينبغي اختزانها وتكشيفها لأغراض الاسترجاع ، ثم إعادة هيكلتها لأغراض العرض. وهناك أسلوبان أساسيان مختلفان تمام الاختلاف لاختزان واسترجاع بيانات لغة الترميز المعيارية العامة ولغة الترميز الموسعة من نظم قواعد البيانات، يعرف أولهما

بأسلوب مركزية البيانات data-centric ، ويعرف الثاني بأسلوب مركزية الوثائق document-centric .

ففي أسلوب مركزية البيانات يُنظر إلى الترميز بلغة الترميز المعيارية العامة ولغة الترميز الموسعة على أنه تركيبة نقل transport syntax في المقام الأول ، وفي هذا المفهوم يوفر نظام قاعدة البيانات أدوات لاستيراد وتصدير بيانات مرمزة ، إلا أنه يختزن عناصر البيانات نفسها في تركيبات علاقية ، أو في تركيبات كيانية object-oriented ، وعادة ما يشيع استخدام نظم مركزية البيانات هذه لاختزان البيانات التجارية أو العلمية التي يمكن تبادلها في صيغتي لغة الترميز المعيارية العامة ولغة الترميز الموسعة ، والتي لا يعتمد ترميزها في أية استخدامات أخرى.

أما نظم مركزية الوثائق فتُستخدم بوجه عام عندما يكون الترميز جزءاً لا يتجزأ من المحتوى نفسه، كما هو الحال مع نظم نشر الكتب والمقالات. فهذه النظم عموماً تستخدم قواعد بيانات أصلية في صيغتي لغة الترميز المعيارية العامة لغة الترميز الموسعة ، تستطيع بالفعل أن تحتفظ ببنية الوثيقة بكاملها. وهناك نظم إدارة المحتوى التي هي مجموعة من التطبيقات التي تجمع كل الوظائف الخاصة بمفهوم مركزية الوثائق ، والتي تشمل عمليات التجميع، والتخزين والصيانة ، والبحث ، والعرض .

وتجدر الإشارة إلى أن إدارة ما وراء البيانات يمكن أن تتم بأي من هذين المفهومين؛ أي أنه لا حرج في اعتبار الترميز بلغة الترميز المعيارية العامة ولغة الترميز الموسعة بمثابة تركيبة ناقلة فقط لتسجيلات ما وراء البيانات ، ومن ثم اختزان واسترجاع عناصر ما وراء البيانات من نظام

إدارة بيانات تقليدي كقواعد البيانات العلاقية، ولا حرج أيضاً في اعتبار تسجيلات ما وراء البيانات وثائق مستقلة يمكن معالجتها بشكل أفضل عن طريق نظم إدارة المحتوى، وهذا يتناسب تماماً مع الخطط الأكثر تعقيدًا المعتمدة على لغة الترميز المعيارية العامة ولغة الترميز الموسعة SGML/XML-based schemes، مثل الوصف الأرشيفي المرمز EAD الذي تكون فيه لما وراء البيانات العديد من خصائص المستندات النصية.

وقد بدأ وكلاء النظم المتكاملة للمكتبات إضافة إمكانيات إنشاء وتكشيف وعرض خطط ما وراء البيانات القائمة على لغة الترميز المعيارية العامة لترميز النصوص ولغة الترميز الموسعة إلى أنظمتهم ، بل وفروا ما يسمح للمكتبيين بتعريف خططهم المحلية الخاصة، ومع أن هذه النظم الوظيفية الفرعية modules [التي يطمح بإضافتها] ليست بالتطور الذي عليه النظام الوظيفي الفرعي الخاص بمارك ، فإنها على أية حال نقطة ستركز عليها اتجاهات التطوير المستقبلي، ورغم أنها ليست إلا مجرد منتجات إضافية add-on products متوافرة بتكلفة إضافية حتى وقتنا الحالي، فإننا نتوقع دمجها مع النظم الأساسية، بل الأكثر من ذلك يمكن القول إن هناك توجهاً بات على وشك التحقق لدمج إمكانات معالجة مارك وخطط ما وراء البيانات الأخرى في غير صيغة مارك.

READINGS :

- Furrie, Betty. Understanding MARC Bibliographic: Machine-Readable Cataloging. ٥th ed., ١٩٩٨.

  Available at http://www.loc.gov/marciumb/.

A simple introduction to MARC basics written for the reader with no prior knowledge of MARC.

- MARC Standards (page of the Library of Congress Network Development and MARC Standards Office website).

  Available at http://lcweb.loc.gov/marc/.

  This site contains extensive, authoritative documentation related to MARC, including links to the MARC٢١ Concise Format documents and code lists, information about Proposed and recent changes to MARC٢١, and representations in SGML and XML.

- Miller, Eric. "An Introduction to the Resource Description Framework.» D-Lib Magazine (May ١٩٩٨).

  Available at : http://www.dlib.orgidlib/may٩٨/miller/٠٥miller.html.

  An early but easily understandable introduction to RDF. For more complete and more recent documentation, see the website of the W٣C's RDF activity at http://www.w٣.orgiRDF/.

- Text Encoding Initiative Consortium. A Gentle Introduction to XML (extracted from TEl P٤: Guidelines for Electronic Text Encoding and Interchange, ٢٠٠٢).

Pdf Available at. http://www.teic .orgi Guidelines٢/ gentleintro.

الفصل الثالث

اللغات المقيدة وخطط التصنيف والمعرفات

أشرنا من قبل إلى أن خطط ما وراء البيانات عادة ما تتألف من الجوانب الدلالية أو الدلالات semantics ، والتركيبة syntax ، وقواعد المحتوى التي تعنى بضبط الكيفية التي تسجل بها قيم عناصر ما وراء البيانات، وتبين كيفية تحديد قيمة كل عنصر من هذه العناصر (وذلك أشبه بما يحدث مع قواعد تحديد المدخل الرئيسي في قواعد الأنجلو أمريكية للفهرسة)، والصيغة التي يتم بها تمثيل هذه القيمة (كما في إرشادات معيار الأيزو رقم ٨٦٠١ لتمثيل التواريخ ISO ٨٦٠١ standard for representing dates)، وتبين كذلك مجموعة أو متوسط القيم التي يمكن أن تعطى لأي من هذا العناصر (مثال ذلك ما يتم في القائمة الاستنادية). ونتناول في هذا الفصل ثلاثة أنواع خاصة من الأدوات التي تتصل بقواعد المحتوى هي: اللغات المقيدة، وخطط التصنيف، والمعرفات.

المفردات أو اللغات المقيدة **vocabularies** :

في سياق اللغة الدارجة ، تمثل مجموعة الكلمات والعبارات التي يستخدمها شخص معين للتعبير عن نفسه لغة هذا الشخص أو مجموعة مفرداته اللغوية، ولا شك أن لكل شخص مفرداته اللغوية أو لغته التي تختلف عن لغات الآخرين؛ فلغة طفل في الثالثة من عمره تختلف اختلافاً واضحاً عن لغة والديه، كما أن لغة عالم الرياضيات تختلف تماماً عن لغة الممرض أو ميكانيكي السيارات أو اختصاصي المكتبات .

أما في سياق ما وراء البيانات، فالمفردة أو اللغة المقيدة يقصد بها عالم

القيم التي يمكن استخدامها لعنصر معين من عناصر ما وراء بيانات، مع ملاحظة أن بعض هذه العناصر - كعنصر العنوان مثلاً - قد يحيط بالمصطلحات المستخدمة للتعبير عنها بعض القيود، هذا في الوقت الذي تكون فيه القيم المحددة لبعض العناصر محددة تحديداً دقيقاً. أما اللغة المقيدة بمعناها الواسع فهي مجموع المحددات التي تحيط بالقيم التي يمكن أن تعطى لعنصر ما وراء البيانات، أما بمعناها الضيق فهي مجموعة محددة مسبقاً للقيم المسموح باستخدامها ، وفي هذا الصدد تنص المجموعة الأساسية لجمعية المصادر المرئية ِVRA Core على أن قيم البيانات الصحيحة الخاصة بعنصر نوع التسجيل هي : مصطلح "عمل work" ، ومصطلح "صورة image "، وهذه المجموعة تعد نمطاً من أنماط اللغة المقيدة البسيطة[1] . أما مكنز الفنون والعمارة the Art and Architecture thesaurus [2]، بحكم ما يتضمنه من مصطلحات متخصصة في مجالي الفنون والعمارة، والتي قد تصل إلى نحو ١٢٥٠٠٠ مصطلح، فيصنف على أنه من فئة اللغات المقيدة الموسعة جداً.

وتعد قوائم المصطلحات والملفات الاستنادية والمكانز هي الطرق الرئيسية التي يستخدمها منشئو ما وراء البيانات لتعيين المصطلحات (سواء أكانت كلمات أو عبارات) من مجموع المفردات اللغوية عند إعداد اللغات المقيدة. حيث تستخدم قوائم المصطلحات البسيطة عندما يكون عدد

---

(١) سبقت الإشارة إلى أن هذه الخطة تخصص عنصراً يعرف " بقيم البيانات " ينص فيه على الاعتماد في اختيار القيم على اللغات المقيدة [المترجم].

(٢) يعد هذا المكنز أحد المطبوعات الهامة التي أصدرها معهد جيتي للمعلومات [المترجم].

المصطلحات صغيراً نسبياً، وعندما تكون الدلالات اللغوية أو معاني هذه المصطلحات واضحة إلى حد كبير، أما الملفات الاستنادية والمكانز فتستخدم في حالة اللغات الأوسع والأكثر تعقيداً، ذلك أن الملفات الاستنادية هي مجموعة مصطلحات معتمدة تستخدمها إحدى الجهات أو تستخدم في سياق قاعدة بيانات معينة، مع ملاحظة أن هذه الملفات الاستنادية لا تعرض العلاقات بين المصطلحات (عدا علاقة الربط بين المصطلحات غير المستخدمة والمصطلحات المستخدمة)[١] ، ولذلك تعد مناسبة جداً أيضاً للغات المقيدة المطلقة flat vocabularies (أي غير الهرمية) ، مثل اللغات الخاصة بأسماء الأشخاص والهيئات. أما المكانز فهي بمثابة تنظيم معين لإحدى اللغات المقيدة بحيث تدرج بموجب هذا التنظيم جميع المصطلحات المعتمدة وإظهار ما بينها من علاقات.

وتتمثل العلاقات التي يظهرها المكنز في علاقات الترادف والتجانس والطبقية أو الهرمية ومدى الارتباط[٢]. ويحدث الترادف عندما يعبر عن مفهوم واحد بمصطلحات مختلفة (بما في ذلك المترادفات والأشكال المختلفة التي تكتب بها الكلمات)[٣]، وفي هذا الحالة عادة ما يتم اختيار مصطلح

---

(١) عادة ما يشار إليها بإحالة أنظر التي تحيل من المصطلح غير المستخدم إلى المصطلح المستخدم [المترجم].

(٢) National Information Standards Organization, NISO Z٣٩,١٩-١٩٩٣(RI٩٩٨), Guidelines for the Construction, Format, and Management of Monolingual Thesauri (Bethesda, Md.: NISO Press, ١٩٩٨), available at http://www.niso.org/standards/resources/Z٣٩-١٩.pdf.Accessed ٦ June ٢٠٠٢.

(٣) كما هو الحال بين كلمة ببليوجرافيا وببليوغرافيا عندما تكتب بالعربية [المترجم].

واحد من هذه المترادفات واعتباره المصطلح الواصف descriptor ، عمل قائمة بالمصطلحات المرادفة تعرف بالمصطلحات البديلة " use for terms " ، ومثال ذلك ما جاء في مكنز الفنون والعمارة ، حيث استخدم مصطلح "مساكن غير المتزوجين single-family dwellings" كمصطلح واصف ، ثم ألحقت به قائمة مطولة بالمصطلحات البديلة غير المستخدمة مثل "منازل غير المتزوجين single-family houses "، و"بيوت غير المتزوجين single-family homes ". أما التجانس فهو عبارة عن سلسلة من الحروف تعطى عندما تجتمع سوياً معاني متعددة[1] ، ولذلك تحرص المكانز على إزالة ما يحيط بالمصطلحات المتجانسة من غموض أو لبس في دلالتها عن طريق إضافة بعض المصطلحات المحددة لها أو المقيدة لاستخداماتها ، مثال ذلك المصطلح التالي الذي ورد هكذا في أحد المكانز[2] Radius (bone) .

أما العلاقة الهرمية أو الطبقية فيُعبر عنها باستخدام "مصطلحات أوسع (BT)" ، و" مصطلحات أضيق (NT)" ، مثال ذلك مصطلحا "ملفات الصور photograph albums" ، و" سجلات الصور والقصاصات scrapbooks" ، اللذان يعاملان على أن كلا منهما مصطلح أضيق من

---

(١) ثمة تفاوت بين العربية والإنجليزية في الجناس ، ففي العربية لدينا نوعان من الجناس : الجناس التام ، والجناس الناقص ، أما في الإنجليزية فثمة ثلاثة أنواع وليس نوعان[المترجم].

(٢) تقصد أن كلمة radius لها معنيان : الأول هو عظم الكعبرة ،أحد عظمي الساعد الأشد قرباً إلى الإبهام ، والثاني هو الشعاع أو نصف القطر ، وعندما تستخدم هكذا : Radius(bone) أي متبوعة بالمحدد (bone) ، فهي تعني تحديد مجال استخدام هذا المصطلح ، وهو في هذه الحالة العظام وليس بمعنى نصف القطر[المترجم].

المصطلح الأوسع "ألبومات albums". ولهذه العلاقة الهرمية أهميتها في عمليات البحث ؛ لأن المواد المكشفة تحت المصطلح الأوسع قد تكون ذات علاقة موضوعية أو مرتبطة بعضها ببعض بالنسبة لبحث أكثر تحديداً ، بمعنى أن المستفيد الذي يبحث تحت مصطلح "سجلات الصور والقصاصات" ولا يجد شيئاً ، يمكنه البحث بمصطلح "ألبومات" . وهنالك كذلك المصطلحات المتصلة أو ذات الصلة Related Terms التي تستخدم عندما يتعدى اهتمام الباحث مفهوماً معيناً إلى مفهوم آخر غيره ذي صلة به ، ومن ثم فالعلاقات الكائنة بين المصطلحات ذات الصلة هي علاقات تبادلية reciprocal (بمعنى أنه إذا كان المصطلح (أ) ذا علاقة بالمصطلح (ب) ، فبالتبعية يكون المصطلح (ب) ذا علاقة بالمصطلح (أ)) ، وليست علاقة هرمية وإلا سيعبر عن طبيعة العلاقة بينهما في شكل علاقة المصطلح الأوسع بالمصطلح الأضيق BT/NT relationships ، ولذلك نجد أن قائمة رؤوس موضوعات مكتبة الكونجرس على سبيل المثال ـ تضم مصطلحات كثيرة ذات علاقة من هذا القبيل، مثال ذلك مصطلحات: "رؤوساء Presidents " ، و "رؤوساء سابقون Ex-Presidents " ، و"كيانات فنية Art Objects " ، و"أنتيكات Antques " ، و"رياضات مائية Aquatic sports " ، و"زوارق Boats وسباق الزوارق Boating " .

وتجدر الإشارة إلى أن المقصد الأساس من استخدام اللغات المقيدة هو رفع معدلات الاسترجاع ؛ ذلك أنه في سياق اللغة الطبيعية يمكن التعبير عن المفهوم الواحد نفسه بعدة كلمات أو جمل مختلفة، حيث يمكن أن يعبر عن " رقائق الحاسب computer chips "، بـ " الرقائق المصغرة microchips "، أو

بـ " الدوائر المتكاملة integrated circuits " ، كما يمكن أن يعبر عن " الحريات المدنية civil liberties " ، بالحقوق المدنية civil rights " ، وتسمية "وقف المباني dedication of buildings " بـ " أوقاف المباني dedications building " ، بل قد تزداد المشكلة تفاقماً عند التعبير الذي تستخدمه فئات مختلفة لكل منها لغته التي يختلف عن لغات الآخرين للمفهوم نفسه ؛ فما تعنيه كلمة "boo-boo" أي خطأ بالنسبة للطفل ، تعني أمراً آخر بالنسبة لوالديه وهو "كدمة" ، بل تعني أمراً ثالثاً بالنسبة للممرضة وهو " التجمع الدموي". ولذلك فعندما تتوافر النصوص الكاملة للوثائق عند التكشيف ، تكون الفرصة مهيأة بشكل جيد لإظهار المترادفات والتعبيرات البديلة من داخل النص، أما تسجيلات ما وراء البيانات فهي غير ذلك ؛ حيث ينبغي أن تكون جامعة أو محكمة، ومن ثم ينبغي على منشئيها أن يأخذوا حذرهم عند اختيارهم قيم البيانات ، وإلا سيخطئ الباحث الوصول للمصادر ذات الصلة ببحثه ، ولذلك فإن اللغة المقيدة تحرص على تخصيص مصطلح كشفي واحد لكل مفهوم ، ومن ثم تجميع كل المواد المتشابهة تحت الواصفات نفسها. وقد لا يكون ذلك بالطبع مفيداً إلا إذا كانت لغة الباحث مطابقة للغة المقيدة ، وعليه فاللغة المخصصة لخدمة المهنيين في المجال الطبي ينبغي أن تنتقى مصطلحاتها التي تختلف عن مصطلحات تلك اللغة الموجهة لخدمة المتخصصين في الرياضيات، بل قد يصعب على الباحثين من أهل المهنة الواحدة توقع المصطلح الذي اختير من قبل معدي اللغة المقيدة ، لأن قيود تصميم هذه اللغات قد لا تشجع في بعض الحالات على استخدام اللغة الطبيعية، ولذلك فإن الباحث - مثلاً - عن الأعمال حول موضوع " تربية الكلاب dog

breeding " قد لا يفكر في البحث تحت مصطلح " الكلاب - تربية Breeding- dogs "، وقد لا يكون هناك مصطلح مفضل بشكل واضح داخل مجموعة من المترادفات الطبيعية. يضاف إلى ذلك أنه كلما اتسع حجم جمهور المستفيدين ، قلت احتمالات الإجماع على استخدام مصطلح معين ، وصعب كذلك على القائمين على أمر اللغة المقيدة مجاراة المفاهيم الجديدة وتغيير المصطلحات المتداولة . ومن الأمثلة على ذلك ما ورد في قائمة رؤوس موضوعات مكتبة الكونجرس ، من أن الأعمال التي تتناول موضوع " إصلاح مالية الحملات campaign finance reform " ، ترد تحت مصطلح " "تمويل الحملات campaign funds " ، ذلك أن مفردات قائمة رؤوس الموضوعات لم تجار اهتمام المستفيد بالموضوع الأكثر تحديداً .

وهناك طرق كثيرة لتضييق الهوة بين لغة المستفيد ولغة المكشف ، لعل أولها يتمثل في توفير اللغة المقيدة نفسها للمستفيد، وهذا ما دعا كثيراً من المكتبات إلى توفير النسخ المجلدة من قائمة رؤوس موضوعات مكتبة الكونجرس - المعروفة بالكتب الحمراء ـ في أماكن الفهارس العامة لجمهور المستفيدين ، أما في البيئة العاملة على الخط المباشر ، فتقوم نظم الاسترجاع بعمل روابط لقوائم المصطلحات والمكانز المتاحة على الخط المباشر ، أو بإدماج بحث المكانز وعرضها في المكان المخصص لتكوين إستراتيجية البحث ، ومن أمثلة ذلك ما يقوم به مشروع كولورادو للرقمنة the Colorado Digitization Project(CDP) ، حيث أعدت قوائم

بالمصطلحات المتصلة بكولورادو وبأسماء المؤلفين من كولورادو، على أن تتاح إمكانية بحث هذه القوائم من خلال موقع المشروع على الإنترنت، وما على الباحث إلا أن يقوم بنسخ المصطلحات من تلك القوائم عند عملية البحث ولصقها في أماكن البحث بالفهرس الموحد للمشروع(١).

أما الطريقة الثانية فتتمثل في دمج تسجيلات تبين طبيعة العلاقات بين المصطلحات عند عملية الاسترجاع، فالمستفيد الذي يبحث عما يتوافر عن حرب فيتنام في فهرس مكتبتنا(٢) - مثلاً - سيجد الإحالة التالية :

VIETNAM WAR ١٩٦١-١٩٧٥

*search under:

Vietnamese conflict, ١٩٦١-١٩٧٥

حرب فيتنام ١٩٦١ - ١٩٧٥م

ابحث تحت مصطلح:

الصراع الفيتنامي ، ١٩٦١- ١٩٧٥م

وفي هذا الصدد تقوم بعض النظم بإعادة توجيه بحث المستفيد تلقائياً، مع أن هذا التوجه أثار قدراً من الخلاف في وجهات النظر مفاده أن رأياً يرى

(١) William A. Garrison, "Retrieval Issues for the Colorado Digitization Project's Heritage Database," D-Lib Magazine ٧, no. ١٠ (October ٢٠٠١), available at http://www.dlib.org/dlib/ octoberOl/garrison/l٠garrison.html. Accessed ٦ June ٢٠٠٢.

(٢) بالطبع مكتبة جامعة فلوريدا التي تعمل بها المؤلفة [المترجم].

في ذلك توفيراً لإحدى الخطوات، ورأياً آخر يرى في ذلك إرباكاً للمستفيد من جراء ما يحصل عليه من نتائج دون علاقات واضحة بينها وبين مصطلح البحث.

وتعتمد الطريقة الثالثة على الباحثين الذين يجدون بعض المواد ذات الصلة عند إجراء أي بحث ، حيث يمكنهم بعد ذلك فحص تسجيلة ما وراء البيانات لمعرفة المصطلحات التي قد خصصت لتستخدم مع الحقول التي تستخدم فيها اللغات المقيدة ، ثم إجراء عملية بحث لاحقة بتلك المصطلحات لاسترجاع المواد ذات الصلة بها بشكل أكبر ، وتسعى بعض النظم الأكثر صلة إلى تيسير هذه العملية عن طريق معاملة محتوى تلك الحقول على أنها روابط فائقة تستخدم تلقائياً كوسائل بحث بمجرد النقر عليها.

ومما تجدر الإشارة إليه أن معدلات جدوى هذه الطرق تنخفض الآن في ظل البيئة الشبكية، على عكس بيئة المطبوعات حيث يكون المستفيد مجبراً على إجراء البحث في ملف واحد في المرة الواحدة مع إحاطته دائماً بقواعد البيانات التي يبحث فيها (سواء أكان ذلك في الفهرس البطاقي أو حتى في كشاف الكتاب أو غير ذلك)، وتكون الفرصة مهيأة له ليتآلف مع مفردات أو لغة تلك المصادر، أما الآن فالفرصة مهيأة أمام المستفيد الجالس إلى شاشة الحاسب لينتقي من بين المئات من قواعد البيانات ، بل غالباً ما تتاح له الفرصة لإجراء بحثه عبر عدة قواعد بيانات مختلفة في آن واحد ، ومن ثم سيواجه العديد من اللغات المقيدة، ومن غير المحتمل أن يتآلف مع تلك اللغات جميعها، يضاف إلى ذلك أن بيئة الإنترنت أو الوب بيئة عالمية

بالفعل ومن ثم تتفاقم معها مشكلة تعدد المفردات بتعدد اللغات المتداولة. ولقد أثارت هذه الإشكالات جهود البحث والتطوير للتنقيب عن أساليب تهدف إلى توافق عمل اللغات المقيدة vocabulary Interoperability. ولعل من أبرزها أسلوب المقابلة أو الربط mapping بين المفردات عن طريق تقديم مفردات مدخلية entry vocabularies (كالكشافات) للغات المقيدة ، واستخدام نظم تصنيف أو أنواع أخرى من أنواع الربط المفاهيمي كنقاط مدخلية لعملية الاسترجاع المبنية على المصطلحات term-based retrieval.

ومع ما يحيط باستخدام اللغات المقيدة في عمليات البحث من معوقات ؛ فثمة اعتراف بأن استخدامها في نظم الاسترجاع يرفع من معدلات التحقيق والاستدعاء ؛ ولذلك هناك تشجيع كبير من خطط ما وراء البيانات على استخدام اللغات المقيدة ، بل إن بعضها يجعل استخدامها إجبارياً على الأقل لقيم بعض عناصر ما وراء البيانات. ولكن يجب بشكل عام إيجاد آلية ما لتحديد اللغات المستخدمة، والمصطلحات المأخوذة منها ، حيث نجد- على سبيل المثال - في خطة الوصف الأرشيفي المرمز EAD - أن التيجان التي يعتمد في صياغة قيمها على لغات مقيدة تستخدم التاج المسمى "المصدر source" ، للإشارة إلى هذه اللغات هكذا على النحو التالي:

```
< controlacces>
<subject sources= "lcsh"> Fishery law and legislation – Minnesota.
</subject>
</controlaccess>
```

التصنيف :

تعد خطط التصنيف شكلاً آخر من أشكال اللغات المقيدة، وتعنى هذه الخطط بتجميع المصادر المتصلة بعضها ببعض في بنية هرمية أو في شكل شجري، مع تخصيص رمز notation لكل فرع من فروع هذه الشجرة، وقد يكون هذا الرمز حرفاً ، أو رقماً ، أو حرفاً مع رقم ، كما قد يلحق بالرمز شروح نصية أو واصفات أو كلاهما معاً بهدف توضيح مجال استخدامه وبيان موقع الفرع في الترتيب الهرمي بشكل عام.

وقد تكون خطط التصنيف عامة (أي تطبق على جميع مجالات المعرفة)، أو متخصصة في مجال محدد، أو موضوع معين، أو إنتاج فكري وطني محدد، أو غيرها من أوجه التغطية . وفي جميع الحالات تبدأ خطة التصنيف بالتقسيمات العامة للمجال، ثم تقسمها بعد ذلك إلى فروعها ، ثم فروع الفروع . فتصنيف المكتبة الوطنية الطبية National Library of Medicine يقسم إلى قسمين رئيسيين: " العلوم ما قبل الإكلينيكية preclinical sciences "، والعلوم الطبية medicine sciences "، ثم يفرع كل منهما إلى ثمانية أقسام فرعية، ثم يقسم كل قسم فرعي إلى ثلاث وثلاثين فئة فرعية ، ثم تفرع كل فئة فرعية إلى فرعيات أخرى حسبما يتفق وطبيعة المجالات . ومن المعروف أنه في أي خطة تصنيف يجب تقسيم الفئة إلى فئات فرعية حسب بعض الخصائص أو مبادئ التقسيم principle of subdivision ، حيث يمكن أن تقسيم الأعمال الفنية - على سبيل المثال - بناءً على الأشكال الأدبية، أو حسب الفنان، أو حسب الفترة الزمنية، على أن تحديد خصائص

فئات المستوى الأعلى هو الذي يتقرر بناءً عليه إذا ما كانت جميع أعمال مايكل أنجلو - مثلاً - تقع جميعها في مستوى واحد بعضها مع بعض ، أم أنها ستتوزع حسب نوعيات اللوحات والتماثيل الأخرى وهكذا.

وثمة نوعان أساسيان لنظم التصنيف، هما: التصنيف الحصري enumerative، والتصنيف الوجهي faceted ، ويقوم نظام التصنيف الحصري بإعداد قوائم بجميع الموضوعات المحتملة مع ترميزها في تسلسل هرمي، بحيث لا يصنف العمل إلا في موقع واحد من التسلسل الهرمي. وعلى هذا الأساس يعد تصنيف مكتبة الكونجرس LCC وتصنيف ديوي العشري DDC نظامان حصريان، وكذلك الحال بالنسبة لمعظم نظم التصنيف[١] التي ظهرت في القرن التاسع عشر [الميلادي]. ويوضح الشكل رقم (٣-١) قسماً من تصنيف مكتبة الكونجرس LCC للفئة الفرعية الخاصة بعلم الحيوان التي تحمل الرمز QL ، مع ملاحظة أن الأعمال العامة عن الفراشات تصنف تحت رقم QL٥٤٤.

الشكل رقم (١/٣)

ملخص للفئة الفرعيةَ QL الخاص بعلم الحيوان في تصنيف مكتبة الكونجرس

Subclass QL

QL ١-٩٩١	Zoology
QL ١-٣٥٥	general

(١) تقصد بالطبع نظم التصنيف الببليوجرافي وليس جهود تصنيف المعرفة فقد شهد التاريخ جهوداً كثيرة في هذا الصدد، وخاصة جهود علماء المسلمين عندما كان الغرب يغط في سبات عميق [المترجم].

	Including geographical distribution
QL ٣٦٠-٥٩٩,٨٢	invertebrates
QL ٤٦١-٥٩٩,٨٢	insects
QL ٦٠٥-٧٣٩,٨	chordates. Vertebrates
QL ٦١٤-٦٣٩,٨	Fishes
QL ٦٤٠-٦٦٩,٣	Reptiles and amphibians
QL ٦٧١-٦٩٩	Birds
QL ٧٠٠-٧٣٩,٨	Mammals
QL ٧٥٠-٧٩٥	Animal Behavior
QL ٧٩١-٧٩٥	Stories and anecdotes
QL ٧٩٩-٧٩٩,٥	Morphology
QL ٨٠١-٩٨٠,٩	Anatomy
QL ٩٥١-٩٩١	Embryology

أمـا نظـام التـصنيف الوجهـي فـيقوم علـى تحديـد الخـصائص العامـة العريـضة للمجـالات الموضـوعية ، التـي تـسمى الأوجـه facets، علـى أن يحدد المصنف جميع الأوجه المناسبة لطبيعة العمل داخل كل فئة، ثم يجمع بينها بطريقة مسبقة التحديد. وعلى هذا الأساس يعتبر التصنيف العشري العالمي UDC المستخدم بكثرة في أوربا نظام تصنيف وجهي.

ومن الممكن القول إن خطط التصنيف تقوم في عالم المكتبات بمهمة مزدوجة؛ حيث تعمل وسيلة لتنظيم المعرفة من ناحية، ووسيلة لترتيب الكتب على الرفوف من ناحية أخرى. كما أن رموز التصنيف تعد أساساً تبنى عليها أرقام الطلب التي تدل على مواقع مواد المكتبة على الرفوف، ومع ذلك ينبغي ألا نخلط بين رموز التصنيف وأرقام الطلب، فرموز التصنيف تحدد مكان العمل في التسلسل الهرمي للمعرفة ، بينما تحدد أرقام الطلب هوية المواد على نحو فريد، كما أن خطط أرقام الطلب - وعلى عكس خطط التصنيف - يجب أن توفر لها طريقة معينة لتمييز كل عمل

على حدة في حالة الأعمال المتعددة في الموضوع نفسه ، وكذا في حالة النسخ المتعددة للعمل الواحد نفسه.

وبرغم ما يمكن أن يتحقق من أوجه إفادة كبيرة لخطط التصنيف في سياق البيئة الإلكترونية، فإن ذلك لم يستثمر أو لم يحظ بالاهتمام الكاف، وقد كانت بوابات الإنترنت الرئيسة مثل " ياهو" تستخدم التصنيف في تنظيم المصادر الإلكترونية، كما أن النظم التي تتميز بالقدرة على توفير بحث مبني على خطط التصنيف وتقديم نتائج البحث في ترتيب مصنف تجعل المستفيد وكأنه يتصفح رفاً افتراضياً virtual shelf . كما أن إمكانية تصفح نظام التصنيف نفسه قد يقدم نقطة مدخلية entry point يمكن استخدامها لأغراض البحث ، كما يسهم في معرفة السياق الذي تستخدم فيه مصطلحات البحث ، يضاف إلى ذلك أن رمز التصنيف، بحكم حياديته اللغوية language-neutral ، يمكن أن يوفر آلية تقارب bridging mechanisim بين المفردات المختلفة بل حتى بين اللغات المختلفة، ولكن لسوء الحظ ، نلحظ قلة من النظم هي التي تتيح المكانز المباشرة لأغراض البحث ، بل أقل منها تلك النظم التي تتيح خطط التصنيف المباشرة ، وربما يرجع السبب في ذلك إلى قلة الخطط المتوافرة حتى الآن ، ومع ذلك فهناك بارقة أمل تدعو إلى تحسن هذا الوضع بتوفير كل من تصنيف ديوي العشري DDC ، وتصنيف مكتبة الكونجرس LCC في أشكال مقروءة آليًا[١].

المعرفات :

---

(١) Diane Vizine-Goetz, Using Library Classification Schemes for Internet Resources, available at http://www.oclc.org/ocldman/colloq/v-g.htm.Accessed ١٠ June ٢٠٠٢.

المعرفات شكل آخر من أشكال ما وراء البيانات، كما أن المعرف الببليوجرافي هو سلسلة من التمثيلات أو المحارف تهدف إلى تعريف كيان ببليوجرافي منطقي تعريفاً متفرداً أو مميزاً ، وإذا كانت بعض أنواع المعرفات تُستنبط من واقع البيانات الببليوجرافية ، فإن بعضها الآخر - وإن كان هو الغالب - يخصص عن طريق هيئة، تعرف بهيئة التسمية naming authority ؛ ذلك أن وجود هذه الهيئة يضمن تفرد المعرفات المخصصة في حدود نظام المُعرِّف identifier system (الذي يعرف كذلك بالاسم المميز أو الفريد namespace) .

وتجدر الإشارة إلى أن هناك تفاوتاً واضحاً بين حدود نظم المعرفات ، بمعنى أن هذه النظم تختلف باختلاف المواد التي يمكن تحديدها، وأنواع الكيانات التي تنطبق عليها. ففي الوقت الذي تركز فيه بعض المعرفات على الأعمال ، تركز فيه أخرى على وسائط العرض manifestation ، أو على وحدات العمل items ؛ فعلى سبيل المثال تعد الأكواد العمودية أو ما يعرف بالباركود barcodes ، والترقيمات الدولية الموحدة للكتب ISBNs نوعين من المعرفات، غير أن الجهة التي توكل إليها مهمة عمل الترميز العمودي هي مكتبة التكويد barcoding library ، بحيث يقوم كل كود عمودي بتعريف مادة وحيدة من تلك المجموعة التي تضمها تعريفاً فريداً ، أما بالنسبة للترقيم الدولي الموحد للكتب ISBN ، فتضطلع بمهمته بالنسبة للمطبوعات الأمريكية وكالة الترقيم الدولي الأمريكية(١) U.S ISBN Agency ، بحيث يعين رقم موحد لكل طبعة من طبعات العمل المنشور في شكل معين (

(١) شركة بوكر R.R. Bowker هي وكالة ردمك الوطنية الأمريكية الآن [المترجم].

والتي تعامل على أنها وسيط عرض) تعريفاً فريداً، وبشكل عام يمكن القول إنه بقدر ما يضيق مجال جهة التسمية ، يقل مستوى الكيان الذي تطبق عليه المعرفات.

وثمة عدة أنواع من المعرفات تتداولها النظم الببليوجرافية بكثرة، من هذه المعرفات: الترقيم الدولي الموحد للكتب ، الذي يطبق على جميع المطبوعات أحادية الموضوع أو ما يعرف بالمنفردات وبجميع الأشكال التي تصدر فيها، وهذا الترقيم عبارة عن سلسلة مكونة من عشر تمثيلات[1] مقسمة إلى أربع شرائح يفصل بينها بواصلة ، هكذا: ISBN ٩٠-٧٠٠٠٢-٠٤-٣ ، وتضطلع وكالة الترقيم الدولي الموحد للكتب بتنسيق أعمال هيئات الترقيم الوطنية والإقليمية التي تعنى بتخصيص بدايات الترقيمات الدولية الموحدة ISBN prefix للناشرين [كل في حدود منطقته الجغرافية]، ثم تترك المسؤولية للناشرين أنفسهم بعد ذلك لتخصيص الأرقام الدولية الموحدة لمطبوعاتهم ، مع مراعاة وجوب تخصيص أرقام مميزة للطبعات المجلدة وأخرى للطبعات الورقية ، وثالثة للأشكال الإلكترونية المتنوعة (كما هو الحال بالنسبة للأعمال التي تصدر في صيغتي RocketBook ، و PDF).

كذلك هنالك الترقيم الدولي الموحد للدوريات international Standard Serials Numbers (ISSN) الذي يخصص للمطبوعات المسلسلة ، وهو معرف مكون من ثماني تمثيلات ، تخصص السبعة الأولى منها لتعريف عنوان المطبوع المسلسل ، في حين تخصص الثامنة كرقم ضبط ، وقد

(١) تغير الآن إلى ١٣ رقماً [المترجم].

تكون رقماً أو خانة مشغولة بالحرف " X " ، وعادة ما يعرض الرقم الدولي الموحد للدوريات دائماً في شريحتين من التمثيلات ، كل منها مكون من أربع تمثيلات يفصل بينها بواصلة هكذا: ISSN ١١٤٠-٣٨٥٣ ، مع ملاحظة أن الترقيم الدولي الموحد للدوريات يبنى وفق نظام تخصيص لامركزي مشابه لنظام الترقيم الدولي الموحد للكتب.

وتجدر الإشارة إلى أن كلا الترقيمين - الترقيم الدولي الموحد للكتب والترقيم الدولي الموحد للدوريات - اعتمدا معيارين دوليين من قبل منظمة التقييس الدولية (أيزو ISO) ، ومعيارين وطنيين من قبل معهد المعايير الوطنية الأمريكية، ومنظمة معايير المعلومات الوطنية ANSI/NISO .

وهنالك كذلك معرف مكونات الكتاب وأجزائه Book item and contribution identifier (BICI) ، ومعرف مكونات الدورية وأجزائها Serial item and contribution identifier (SICI)، وهما معرفان وطنيان قياسيان يطبقان على العناصر المكونة للكتب والدوريات ، كالفصول [بالنسبة للكتب]، والأعداد وما تشمل من مقالات [بالنسبة للدوريات]. مع ملاحظة أن المعرف الأول يبنى على نظام الترقيم الدولي الموحد للكتب ، في حين يبنى الثاني على نظام الترقيم الدولي الموحد للدوريات ، وعلى عكس المعرفات السابقة ، لا يقوم هذان المعرفان على عملية التخصيص والتسجيل ، ولكنهما يستنبطان من البيانات الببليوجرافية للمواد المعرفة، حيث يعرف مقال لنيخيل هيوثيننج Linkhil Hutheening المسمى "الحفاظ على دفء المقاعد keeping the Seats Warm "، المنشور بمجلة Forbes ، في ص ٦٢ من المجلد رقم ١٥٧، العدد الأول، الصادر في الأول من يناير ١٩٩٦م، بناء

على نظام معرف مكونات الدورية وأجزائها على النحو التالي :

SICI: ٠٠١٥-٦٩١٤ (١٩٩٦٠١٠١) ١٥٧:١<٦٢:KTSW>٢,٠.TX;٢-F

ويلاحظ في هذا المثال أن الترقيم الدولي الموحد ٠٠١٥-٦٩١٤ يحدد هوية عنوان المجلة ، أما الرقم المحصور بين القوسين (١٩٩٦٠١٠١) فهو مستنبط من تاريخ صدور العدد[١] ، أما "١٥٧:١" فمستنبط كذلك من التسمية الرقمية للمجلد والعدد ، أما "<٦٢:KTSW>" فالشق الأول منه مستنبط من رقم صفحة، أما الشق الثاني فمن الحروف الأولى لكلمات العنوان ، ويأتي بقية المعرف (٢,٠.TX;٢-F) وهو عبارة عن سلسلة من القيم المرمزة الإجبارية التي تستخدم لأغراض محددة كأن تستخدم مجموع تدقيق checksum[٢]، وتعبر عن الصيغة التي تصدر فيها المقالة.

وإذا كان من الممكن تخصيص الترقيم الدولي الموحد للكتب، والترقيم الدولي الموحد للدوريات ، ومعرف مكونات الكتاب وأجزائه ، ومعرف مكونات الدورية وأجزائها للمطبوعات التي تنشر في الصيغتين المطبوعة والإلكترونية، فلدينا كذلك معرف الكيان الرقمي the digital object identifier "DOI" المعني أساساً بالمنشورات الرقمية دون غيرها ، ويتكون هذا المعرف من جزأين: بادئة prefix ، ولاحقة suffix ، أما البادئة فيخصصها نظام لامركزي لهيئات التسجيل التي لا تتميز حسب توزيعها

---

(١) أي الأول من يناير عام ١٩٩٦ [المترجم].

(٢) قيمة محسوبة اعتماداً على محتويات مجموعة بيانات وذلك لغرض تدقيق البيانات والكشف عن مدى صحتها [المترجم].

الجغرافي ـ كما هو الحال مع نظامي الترقيم الدولي الموحد للكتب والترقيم الدولي الموحد للدوريات، وإنما بناءً على طبيعة مجتمع المستفيدين الذي تخدمه ونوعية الخدمات التي تقدمها له. وعادة ما تستهل البادئة دائماً بالرقم "١٠." ، ثم يتبعه تعريف بهوية منتج المحتوى الذي سجل معرف الكيان الرقمي ، أما اللاحقة فهي سلسلة عشوائية من الأرقام تحدد هوية المحتوى الفعلي للمادة موضع التعريف، فعلى سبيل المثال ، يكون معرف الكيان الرقمي للنسخ المنشورة على الإنترنت في صيغة لغة ترميز النص الفائق HTML من الدليل الإرشادي لمعرفات الكيان الرقمي DOI Handbook ، هكذا: "١٠,١٠٠٠/١٠٢" ، حيث يلاحظ أن السابقة "١٠٠٠ " تعني أن معرف الكيان الرقمي قد خصص من قبل مؤسسة الدولية لتعريف الكيانات الرقمية the International DOI foundation (IDF) ، بينما تحدد اللاحقة "١٠٢ " طبيعة العمل بالتحديد، وعليه فإن معرف الكيان الرقمي للنسخ المنشورة من هذا العمل نفسه ولكن في صيغة PDF يكون هكذا : "١٠,١٠٠٠/١٠٦".

ولعل صفة التفاعلية actionable التي يتسم بها معرف الكيان الرقمي هي التي تدفع ناشري المحتوى الرقمي لتداوله على نطاق واسع ، ويقصد بتفاعلية هذا المعرف أنه يمكن أن يوظف كرابط سريع مباشر عند النقر عليه clickable hotlink في المتصفح، ومن ثم يمكن تحويله أو ترجمته إلى المحدد الموحد لمكان المصدر URL ، عن طريق نظام التحويل Resolution system الذي تقوم به المؤسسة الدولية لتعريف الكيانات الرقمية the International DOI foundation (IDF) . ونتيجة لذلك فقد بات معرف الكيان الرقمي متداولاً في عملية الربط المرجعي أو بين المراجع، وما شابه ذلك

من عمليات؛ حيث يمكن لإحدى المقالات أن تستشهد - مثلاً - بمقالة أخرى منشورة قبلها مكتفية في الإشارة المرجعية بإعطاء معرف الكيان الرقمي لتلك المقالة السابقة، وعندما ينقر القارئ على ذلك المعرف يوجه المعرف إلى مترجم أو محلل المؤسسة الدولية لتعريف الكيانات الرقمية IDF's DOI resolver ، الذي يقوم بدوره بفحص المعرف ومن ثم استرجاع المحدد الموحد لمكان المصدر URL المرتبط به ، ثم بعد ذلك يعيد توجيه الطلب إلى وجهة المحدد الموحد لموقع المصدر URL المستهدف .

ومما يلفت الانتباه أن ليس هناك حتى الآن سبب وجيه لعدم وجود نظام مشابه لنظام التحويل أو الترجمة للأنواع الأخرى من المعرفات(١) ، إلا أن الاتحاد الأوربي في الواقع يعمل الآن على اختبار نظام تحويل معرف مكونات الدورية وأجزائها SICI ، ومع ذلك فقد كانت البنية التحتية لدعم عملية التحويل دائماً جزءاً من نظام معرف الكيان الرقمي DOI system.

وإذا ما طبق المفهوم السابق على المحدد الموحد لمكان المصدر لا نستطيع أن نعده معرفًا؛ ذلك لأنه يحدد مكان المادة الفعلية ولا يعرف بهوية الكيان المنطقي، فالمحدد الموحد لموقع المصدر يحدد خدمة الوصول (التي عادة ما تكون خدمة بروتوكول نقل النص الفائق HTTP ، أو خدمة بروتوكول نقل الملفات FTP ، أو تلنت TELNET أو غيرها) ، ويحدد كذلك مكان المادة داخل تلك الخدمة. ولأن نقل الكيانات الرقمية من خادم إلى آخر يعد من الأمور الشائعة ، فقد يكون للكيان الواحد أكثر من محدد موحد واحد

(١) تقصد المعرفات الأخرى غير معرف الكيان الرقمي [المترجم].

URLs علـى مـدار حـياته ، كمـا سـيكون للمحـتوى نفسه محـددات مـوحدة URLs مختلفة باختلاف الخدمات التي تتيحه.

وقد أقرت لجنة العمل الخاصـة بهندسـة الإنترنت IETF - تلك المنظمة التـي تشرف علـى تطوير معايير الإنترنت - لفترة مـا بأوجه القصور التي تعتري المحدد الموحد لموقع المصدر، والحاجة إلى دعم مجموعة من المعرفات الدقيقة في البيئة الشبكية ؛ وتلبية لهذه الحاجة طورت مجموعات عمل تابعة لهذه اللجنة إطار الاسم الموحد لموقع المصدر URN framework ، بحيث إن هذا الاسم الموحد لموقع المصدر يمكن - كما هو الحال مع معرف الكيان الرقمي - تحويله إلى محدد موحد واحد لموقع المصدرURL ، أو أكثر من محدد URLs ، كما أنه يتكون من جزءين هما: سلسلة من التمثيلات تحدد هيئة التسمية، وسلسلة أخرى تخصصها هيئة التسمية لتحديد هوية الكيان المعرف.

وتجدر الإشارة إلى أن أنواعاً أخرى من المعرفات يمكن أن تعمل داخل إطار الاسم الموحد للمصدر مادام في الإمكان التعبير عنها بشكل مناسب في تركيبة هذا الاسم ، حيث يمكن على سبيل المثال التعبير عن الترقيم الدولي الموحد للكتاب كاسم موحد للمصدر هكذا : "URN:ISBN:٠-٣٩٥-٣٦٣٤١-١" ، مع ملاحظة ضرورة وجود السابقة URN لتعريف هذه السلسلة من التمثيلات على أنها اسم موحد للمصدر، كذلك ضرورة وجود رمز الترقيم الدولي الموحد للكتاب - أي "ISBN" - لتحديد هوية هيئة التسمية ، أما الرقم "٠-٣٩٥-٣٦٣٤١-١" فهو معرف الكيان الرقمي الذي خصصته هيئة الترقيم الدولي الموحد للكتب.

وينبغي أن يكون هنالك داخل إطار الاسم الموحد للمصدر آلية تتسم بالكونية أو العالمية global إلى حد كبير لتوجيه الأسماء الموحدة إلى خدمات التحويل المناسبة لهيئة التسمية المعنية ، إلا أن آلية من هذا النوع لم يتحقق وجودها بعد ، كما لم يتحقق وجود متصفح أصلي داعم لتلك الأسماء الموحدة على عكس ما هو قائم بالنسبة للمحددات الموحدة ، ولذلك تشير العديد من خطط ما وراء البيانات التي تتطلب وجود معرفات تفاعلية لعناصر بيانات محددة إلى استخدام معرفات موحدة لمواقع المصادر "uniform Resource Identifiers" URIs.

أما في سياق إطار لجنة العمل الخاصة بهندسة الإنترنت ، فتعامل كل من المحددات الموحدة لمواقع المصادر URLs ، والأسماء الموحدة للمصادر URN كمعرفات موحدة للمصادر P وعليه فإن تحديد المعرفات الموحدة للمصادر كمحتوى لما وراء البيانات يتيح الفرصة لاستخدام المحددات الموحدة لمواقع المصادر على المدى القصير ، واستخدام الأسماء الموحدة للمصادر على المدى البعيد.

وبحكم الدور الحيوي الذي يُنتظر أن تنهض به المعرفات في البيئة الرقمية، فقد تواجدت - فضلاً عن المعرفات المشار إليها في الفقرات السابقة - معرفات أخرى خاصة بأنواع معينة من المواد ، وقد أقرت بعض هذه المعرفات ، ولما يقر بعضها الآخر حتى الآن ، لأنها لا تزال قيد التطوير ، ومن هذه المعرفات : الرقم المعياري الدولي للتقارير الفنية international standard technical report number "ISRN المعني بالتقارير الفنية ، والرمز المعياري الدولي للتسجيلات International standard recording code

"ISRC" لتعريف التسجيلات الصوتية والسمعية المسجلة على الأسطوانات المدمجة CD وغيرها من الوسائط ، والرقم المعياري الدولي للموسيقى International standard music number "ISMN" الخاص بالمطبوعات الموسيقية ، والرمز المعياري الدولي للعمل الموسيقي international standard musical work code "ISWC" ، والرقم المعياري الدولي للوسائط السمعية البصرية international standard Audiovisual Number "ISAN" ، اللذان يعرفان الأعمال المناظرة لمطبوعات أو أساليب عرض معينة ، وأخيراً الكود المعياري الدولي للعمل النصي international standard Textual Work code "ISTC" الذي يجري تطويره لتعريف الأعمال النصية.

READINGS :

- Lynch, Clifford A. "Identifiers and Their Role in Networked Information Applications." ARL: A Bimonthly Newsletter of Research Library Issues and Actions ١٩٤ (October ١٩٩٧). Available at http://www.arl.org/newsltr/١٩٤/identifier.html.

- Metadata Resources (page of the UKOLN Metadata Resources Clearinghouse) has a fairly comprehensive bibliography related to identifiers and URIs. Available at http://www.ukoln.ac.uk/metadata.resources/.

- Milstead, Jessica L. Use of Thesauri in the Full-text Environment. ١٩٩٨. Available at http://www. jelem.corn/useof.htm/. Accessed ١٠ June ٢٠٠٢.

- Sevonius, Elaine. "Use of Classification in Online Retrieval:'

Library Resources and Technical Services ٢٧, no. ١ (January/March ١٩٨٣): ٧٦-٨٠.

- Two papers exploring uses of thesauri and classification, respectively, in the online retrieval environment.

الفصل الرابع

سبل تحقيق التشغيل البيني

تعنى التطبيقات العاملة في البيئة الشبكية بأنواع متعددة من التشغيل البيني أو المتداخل Interoperability ، وعندما يذكر التشغيل البيني على إطلاقه يقصد به أكثر من معنى ؛ منها أن تطبيقين يتقاسمان بروتوكول اتصال مشترك بينهما، ومنها كذلك أن حاسباً عميلاً واحداً client يستطيع التعامل مع العديد من الحاسبات الخادمة servers ، كما يقصد به كذلك إمكانية إعادة استخدام البيانات في سياقات مختلفة . أما عندما نتحدث عن التشغيل البيني في سياق ما وراء البيانات ، فإننا نتحدث عموماً عن عمليات البحث البيني أو المتداخل Search Interoperability ، أي القدرة على إجراء عمليات البحث عبر مجموعات متباينة من تسجيلات ما وراء البيانات ، ومن ثم الحصول على نتائج جيدة ، مع أن ما وراء البيانات قد تنشأ وفق خطة واحدة إلا أن منشئيها - سواء أكانوا أفراداً أو هيئات - يمثلون أطيافاً مختلفة، أو أنها قد تكون نتاج تطبيق خطط متعددة وليس خطة واحدة.

الفهارس الموحدة **Union Catalogs** :

يمكن أن تتحقق عملية البحث البيني بأكثر من طريقة واحدة ، ويعد بناء قاعدة بيانات مركزية لما وراء البيانات إحدى هذه الطرق، كما تعد الفهارس الموحدة للمكتبات المعتمدة في بنائها على صيغة مارك مثالاً جيداً لهذه الطريقة ، وهذه الفهارس الموحدة يمكن أن تتحقق على أي مستوى ، بدءاً بالمستوى المؤسسي (كما هو الحال عندما ينشأ فهرس موحد لمكتبة عامة مركزية بفروعها المتعددة) ، وانتهاءً بالمستوى الدولي (كما هو الحال مع الفهرس العالمي

المعروف بـ OCLC's WorlCat الذي يتبناه مركز الحاسب للمكتبات على الخط المباشر). ومن الملاحظ أن العديد من ائتلافات تقاسم الموارد على مستوى الولاية statewide والإقليمية تدعم فكرة الفهارس الموحدة أو بالأحرى الملفات الموحدة، ومن أمثلة ذلك القوائم الموحدة للمسلسلات.

وثمة نماذج مختلفة يسترشد بها عند إنشاء الفهارس الموحدة ؛ حيث تقوم المكتبات المشاركة في مشروع الفهرس الموحد في أحد هذه النماذج بإرسال نسخة من تسجيلات فهارسها إلى هيئة تضطلع بمهمة إعداد فهرس مركزي قابل للبحث ، ومن أمثلة هذا النموذج مشروع ميلفيل MELVYL ، الفهرس الموحد القديم لجامعة كاليفورنيا الذي يحدِّثه أسبوعيًا أو شهريًا تسعة وعشرون مصدراً مختلفاً من مصادر البيانات ، منها أقسام الفهرسة بمكتبات جامعة كاليفورنيا، ومكتبة ولاية كاليفورنيا، ومكتبات المعاهد الأخرى التابعة للجامعة(١).

وفي سياق نموذج آخر يمكن إنشاء التسجيلات مباشرة في قاعدة بيانات الفهرس الموحد ، ثم يتم بعد ذلك تزويد النظام المحلي للمكتبات المشاركة بنسخ من هذه التسجيلات. وفي أي من هذين النمطين يمكن الاحتفاظ بجميع التسجيلات التي أعدت للعمل الواحد من قبل جميع الجهات المشاركة كنسخ مكررة ، أو يمكن دمجها في تسجيلة رئيسية واحدة ، ولكن يوضح فيها المكتبات التي تقتني نسخاً منها.

وهناك أسلوب ثالث يتم بمقتضاه إنشاء فهرس موحد غير حقيقي

---

(١) Karen Coyle, "The Virtual Union Catalog: A Comparative Study," D-Lib Magazine ٦, no. ٣ (March ٢٠٠٠), available at http://www.dlib.org/dIib/marchOO/coyle/٠٣coyle.himI.Accessed ٢٦ July ٢٠٠٢

pseudo-union catalog عن طريق بناء كشاف موحد من ملفات الفهارس المتعددة بدلاً من إعداد قاعدة بيانات شاملة ، وعند اختيار مداخل معينة من الكشاف ، تعرض تسجيلاتها من الفهارس الأصلية أو المصدرية.

وللفهارس الموحدة من هذا القبيل جدواها الواضحة ، ذلك أن مجتمع المكتبات برمته يشترك في صيغة بيانات مشتركة (ممثلة في مارك٢١) ، كما يشترك بشكل أو بآخر في مجموعة واحدة من قواعد الفهرسة ، يضاف إلى ذلك أن التسجيلات المختزنة في الملف المركزي ليست متجانسة إلى حد ما بعضها مع بعض فحسب، بل إنها متشابهة مع تلك التي تختزن في الفهارس المحلية للمكتبات المشاركة ، ومن ثم يشعر الباحث في الفهرس الموحد بألفة في تعامله مع خدمات البحث والاسترجاع التي يوفرها ذلك الفهرس.

ومع ما يحيط بإنشاء الفهارس الموحدة اعتماداً على مصادر غير متجانسة لما وراء البيانات من تعقيد ، فإن تحقيق ذلك يعد أمراً ممكناً، وإحدى السبل التي يمكن أن تعين في تحقيق ذلك تحويل خطط ما وراء البيانات المتعددة إلى صيغة اختزان وتكشيف موحدة. وقد ثبتت جدوى دبلن كور في هذا الصدد، حيث يمكن أن تكون مجموعة بينها قواسم مشتركة من العناصر التي يمكن أن تضاهي أو تربط mapping بمجموعة عناصر خطط أخرى أغنى منها. ولقد تبنى هذا المفهوم مشروع كولورادو للرقمنة Colorado digitization Project (CDP) الذي يرعى قاعدة موحدة لما وراء البيانات التي شاركت فيها أرشيفات كولورادو وجمعياتها التاريخية ومكتباتها ومتاحفها (http://coloradodigital.coalliance.org)، إذ يسمح لهذه الجهات المشاركة أن ترسل[تسجيلات] ما وراء بيانات بأي صيغة كانت ، شريطة أن يقدم كل منها وصفاً بكيفية مقابلة عناصر ما وراء البيانات

الخاصة بها بما يقابلها من عناصر ما وراء البيانات المعتمدة في دبلن كور، ثم يأتي دور المشروع في قيامه بعملية التحويل المركزي لتسجيلات ما وراء البيانات الخاصة بالجهات المشاركة إلى تلك الصيغة المشتركة أو الموحدة قبيل تحميلها في الفهرس الموحد.

وهناك مثال آخر تقدمه المواقع التي طبقت بروتوكول تجميع ما وراء البيانات لمبادرة الأرشيفات المفتوحة Open Archives Initiative Metadata harvesting protocol . ونموذج الأرشيفات المفتوحة هذا هو شكل مختلف عن الفهرس الموحد التقليدي ، حيث يتم فيه جمع ما وراء البيانات من مختلف المواقع المشاركة عن طريق الجمع harvesting وليس عن طريق المشاركة contribution (أي أنه نموذج سحب وليس نموذج دفع a pull rather than push model) ، وهو يشبه في ذلك الطريقة التي تجمع بها عناكب الإنترنت محتويات الصفحات المهيأة في صيغة لغة ترميز النص الفائق HTML. كما أن بروتوكول جمع ما وراء البيانات نفسه بروتوكول بسيط ، حيث تنحصر مهمته في نقل الاستفسارات والاستجابات إلى بروتوكول نقل الملفات الفائقة HTTP ، أما التطبيق المسمى بالجامع harvester فيطلب من مستودع ما وراء البيانات تزويده بقائمة لصيغ ما وراء البيانات التي يدعمها ، أو يطلب منه قائمة بمجموعات التسجيلات التي يدعمها ، أو أن يطلب منه كذلك قائمة بمعرفات جميع التسجيلات الموجودة في المستودع أو في مجموعة التسجيلات أو بهما معاً ، كما يمكن لهذا الجامع كذلك أن يطلب من المستودع ، أن يقوم بتصدير تسجيلة واحدة لما وراء بيانات أو مجموعة من التسجيلات. وينبغي في هذه الحالة أن يكون مقدم البيانات المتوافقة مع مبادرة الأرشيفات المفتوحة قادراً على الاستجابة لتلك الطلبات وتصدير ما

وراء البيانات في صيغة واحدة على الأقل غالباً ما تكون تمثيل عناصر دبلن كور غير المحدد في صيغة لغة الترميز الموسعة XML ، وهذا لا يعني أن ترفض الصيغ الأخرى ، وإنما يمكن أن تقدم مشروطة بالاتفاق على ذلك بين كل من مقدم البيانات والجامع harvester ، وتجدر الإشارة في هذه الحالة إلى أن عملية التحويل إلى الصيغة المشتركة يقوم بها جميع مقدمي البيانات المشاركين، ولا تتم في الموقع المركزي كما هو الحال مع مشروع كلورادو للرقمنة CDP.

ويمكن كذلك تجميع ما وراء البيانات التي تتضمنها في ملف المصدر الخاص بصفحات الوب المهيأة في لغة ترميز النص الفائق في قاعدة بيانات موحدة، ثم تكشيفها حتى تتمكن محركات بحث الإنترنت من استرجاعها. وكما هو الحال مع جميع أنواع البحث الموحد، فإنه بقدر ما تكون البيانات متناغمة، تكون معدلات الاسترجاع مرتفعة، وفي هذا الصدد طورت عدة ولايات - كولايتي واشنطن وإلينوي - أدلة إرشادية لاستخدام تيجان ما وراء البيانات أو ما تعرف بتيجان الميتا <META> لمشروعات خدمة محدد مواقع المعلومات الحكومية المعروف بجيلس Government Information Locator Service (GILS) على مستوى الولاية ، حيث تقوم زواحف الوب على هذا المستوى state webcrawlers بزيارة مواقع الهيئات الحكومية بحثاً عن تيجان ما وراء بيانات أو تيجان ميتا <META> محددة لأغراض تكشيفها .

ومن الملاحظ في هذا الصدد أن الفهارس الموحدة المتجانسة لا تقوم جميعها بتحويل ما وراء البيانات إلى صيغة مشتركة، ففي نموذج آخر يحتفظ بالتسجيلات الأصلية أو المصدرية لما وراء البيانات في خططها الأصلية وبصيغها التي هي عليها، ولكن يجري بحثها وكأنها ملف واحد.

ويعد مشروع الذاكرة الأمريكية الذي تتبناه مكتبة الكونجرس(http://memory.loc.gov) من أشهر النماذج على هذا الاتجاه؛ حيث يتيح موقع هذا المشروع على الإنترنت أكثر من سبعة ملايين مادة من أكثر من مئة مجموعة مقتنيات، ومع أن الغالبية العظمى من هذه المجموعات كان ضمن مقتنيات مكتبة الكونجرس وحولت إلى مجموعات رقمية ، فقد مولت مؤسسات أخرى لتساهم بمجموعاتها الرقمية في سياق منافسة أميريتك التي تبتنها مكتبة الكونجرس LC Ameritech competition خلال الفترة من عام ١٩٩٧ إلى ١٩٩٩م، حيث سمح للمؤسسات المشاركة باختزان محتوياتها الرقمية محلياً ، على أن تقوم بإرسال ما وراء البيانات في صيغة واحدة من صيغ خطط أي من مارك ، أو دبلن كور، أو ترويسة مبادرة ترميز النص TEI ، أو الوصف الأرشيفي المرمز، إلى مكتبة الكونجرس التي قامت بدورها بالاحتفاظ بملفات مستقلة تمثل الأنواع المختلفة لما وراء البيانات ، بالإضافة إلى إنشاء مجموعات مستقلة من الكشافات لتلك الملفات ، وعندما يُقدِم المستفيد على بحث المجموعة برمتها ، فإنه في واقع الأمر يجد نتائج للبحوث المجمعة في كل كشاف من هذه الكشافات ، كما أن اختيار أي مدخل كشفي للعرض يثمر عن استرجاع تسجيلة كاملة المصدر لما وراء البيانات full-source metadata record ، بعد صياغتها في شكل مناسب للعرض(١).

البحث البيني أو البحث عبر النظم **Cross- System Search** :

(١) Carolyn R. Arms, Access Aids and Interoperability, ١٩٩٧, available at http://memory.loc.gov/ammem/award/docs/interop.html. Accessed ١١ June ٢٠٠٢.

في مفهوم الفهرس الموحد يتم إعداد قاعدة بيانات موحدة (أو كشاف موحد في بعض الحالات) لما وراء البيانات ، ثم يُستخدم نظام بحث واسترجاع مركزي للوصول إلى بيانات هذه القاعدة . أما في مفهوم البحث البيني عبر عدة نظم فيتم اختزان التسجيلات في عدة قواعد بيانات موزعة أو لا مركزية ، على أن تسترجع هذه التسجيلات من خلال إمكانات البحث التي تتيحها برمجية البحث الخاصة بكل قاعدة من هذه القواعد.

ويعد البروتوكول المسمى زد ٣٩,٥٠ ANSI/NISO Z٣٩,٥٠ (رقم ISO ٢٣٩٥٠) بروتوكولاً دولياً معيارياً يتيح الفرصة لأحد نظامين (النظام الأصلي أو العميل) أن يجري بحثاً في نظام آخر (النظام المستهدف أو الخادم)، ثم يتلقى النتائج في صيغة يتمكن من عرضها على المستفيد(١). ذلك أن هذا البروتوكول (Z٣٩,٥٠) يقوم على فتح حوار بين نظامين يسمح بمقتضاه للنظام الأصلي أن يتصل بالنظام المستهدف ويجري فيه بحثاً ، بل يطلب أن تعرض نتائج هذا البحث في صيغة محددة ، بل أكثر من ذلك يمكنه أن يحدد عدد التسجيلات التي تسترجع في المرة الواحدة ، وهكذا. وفوق ذلك كله تكمن العبقرية الحقيقة للمعيار Z٣٩,٥٠ في عدم اكتراثه بضرورة الترجمة من لغة البحث الخاصة بكل نظام إلى لغة النظام الآخر، وإنما يطلب بدلاً من ذلك أن يعبر عن كل عملية بحث بتركيبة موحدة مجردة a common abstract syntax ، لذلك لا يحتاج كل نظام إلا إلى معرفة لغتين

(٢) ANSI/NISO Z٣٩,٥٠-١٩٩٥ Information Retrieval: Application Service Definition and Protocol Specification (Washington, D.C.: NISO Press, ١٩٩٥), available at Http://www.niso.org/standards/resources/Z٣٩-٥٠.pdf. Accessed ١١ June ٢٠٠٢.

اثنتين : لغة البحث الخاصة به ، ولغة البحث الخاصة بالمعيار Z٣٩,٥٠. (ويمكن للمرء أن يتخيل المفهوم الأول وكأنه شخص لا يعرف إلا لغة واحدة، ولكي يتواصل مع جميع البشر عليه أن يتعلم جميع لغاتهم ، أما في مفهوم معيار Z٣٩,٥٠ فلا يتطلب الأمر من الشخص الذي يرغب في التواصل مع جميع البشر سوى معرفة لغتين: لغته الأصلية ولغة الاسبرانتو)[١] ، وفي حالتنا هذه ، يقوم الحاسب العميل المدعم بمعيار زد ٣٩,٥٠ clients بترجمة بحث المستفيد إلى اللغة الموحدة Z٣٩,٥٠ للاستفسارات ، ثم يقوم الحاسب الخادم المدعم بمعيار زد ٣٩,٥٠ Z٣٩,٥٠ servers بترجمة الاستفسار المقدم في اللغة الموحدة إلى تركيبة نظام البحث الخاص بالخادم.

ويكتسب المعيار Z٣٩,٥٠ هذه الاستقلالية اللغوية من جراء استخدامه مجموعة الخصائص attributes set ، أو قوائم خصائص البحث المجردة لأنواع معينة من عمليات البحث، وتعد مجموعة الخصائص المعروفة ب "Bib-١" هي الأكثر تداولاً بالنسبة للبحث الببليوجرافي[٢]، إذ تشتمل هذه المجموعة على ستة أنواع من الخصائص هي خاصية : الاستخدام use، و العلاقة relation ، والموضع position ، والبنية structure ، والبترtruncation ، و الاكتمال completeness . حيث تحدد خاصية " الاستخدام" مجموعة

---

(١) لغة دولية موحدة [المترجم].

(٢) Bib-l Attribute Set, available at http://lcweb.loc.gov/z٣٩٥./agency/defns/bibI.html. Accessed ١١ June ٢٠٠٢. See also Attribute Set Bib-l (Z٣٩,٥٠-١٩٩٥) Semantice, available at fip://fip.loc.gov/pub/z٣٩٥٠/defslbibl.txt.

نقاط الوصول التي يمكن مطابقة أو مضاهاة كلمة البحث بها ، إذ تشير الخاصية "١" منها - مثلاً - إلى اسم شخص ، وتشير الخاصية "٢" إلى اسم هيئة ، بينما تشير الخاصية "٣" إلى عنوان المؤتمر. أما بقية الخصائص فتحدد الكيفية التي ينبغي أن يتخذها البحث؛ حيث تحدد خاصية "الموضع" - على سبيل المثال - الموضع الذي ينبغي أن تتواجد فيه كلمة البحث في سياق الحقل ، فتشير الخاصية "١" في هذا الصدد مثلاً إلى أن كلمة البحث يجب أن تقع في بداية الحقل ، بينما تشير الخاصية "٣" إلى إمكان تواجد كلمة البحث في أي موضع من الحقل.

ويسمح المعيار Z٣٩,٥٠ ببناء واجهة استخدام واحدة لمجموعة من النظم الموزعة غير المتجانسة ، بحيث تستخدم هذه الواجهة في إجراء البحث الموسع broadcast (أو ما يعرف بالبحث الاتحادي أو الفيدرالي federal) ، الذي يتم على نحو متزامن في قواعد بيانات متعددة . ويتوافر الآن في الأسواق عدة بوابات للمعيار Z٣٩,٥٠ ، تتوافق كل منها مع بروتوكول نقل الملفات الفائقة HTTP من ناحية ، كما تقوم مقام [برمجية] عميل بروتوكول client Z٣٩,٥٠ من ناحية أخرى(١). وعندما يستخدم المستفيد متصفحه للاتصال بإحدى البوابات تظهر له قائمة بالخدمات المستهدفة التي يسمح له باستخدامها جميعاً أو اختيار ما يراه مناسباً له منها

(١) See, for example, OCLC’s SiteSearch (http://www.sitesearch.oclc.org/), Blue Angel Technology’s MetaStar Gateway (http://www.blueangeitech.com), and Fertwell-Downing’s Zportal http:/www.fdgroup.co.uk/fdi/zportal/overview.html.

، أما نتائج البحث فغالباً ما تعرض في قوائم حسب النظام المستهدف ، وإن كانت بعض البوابات تحاول دمج النتائج المسترجعة من مختلف النظم المستهدفة وعرضها في قائمة واحدة.

وتجدر الإشارة إلى أن مثل هذه البوابات تحظى باهتمام واسع وخاصة في المواقف التي تتطلع فيها إحدى الجهات إلى تقديم إمكانية البحث البيني أو المتداخل لمجموعة من المصادر مسبقة الاختيار ، وقد قام العديد من الائتلافات بإنشاء فهارس موحدة "افتراضية virtual " مبنية على المعيار Z٣٩,٥٠ تتقاسم بينها قاعدة بيانات موحدة مركزية بديلاً عن الفهارس الموحدة "المادية physical " . كما تستخدم هذه البوابات كذلك لأغراض البحث عبر أنواع متباينة من المصادر الببليوجرافية (مثل فهارس المكتبات وقواعد بيانات خدمات التكشيف والاستخلاص A&I databases)، أو عبر مستودعات أنواع مختلفة من ما وراء البيانات مثل مارك ودبلن كور.

ومع أن الهدف الكامن وراء مثل هذه البوابات يتمثل في إتاحة خدمات البحث المتجانس عبر المصادر غير المتجانسة ، فقد أظهرت الممارسات التطبيقية عدداً من أوجه القصور للتوجهات المعتمدة على المعيار Z٣٩,٥٠، ويمكن القول إن شطراً من مسؤولية هذه الإشكالات تتحمله البرمجية(١)، لأن هذا المعيار يعد بروتوكولاً معقداً ويصعب على جميع النظم تنفيذه ، وقد نتج عن ذلك ضعف التنفيذ المكتمل لهذا البروتوكول وظهور التطبيقات الجزئية غير المكتملة له، يضاف إلى ذلك توقع نوع من الاسترجاع غير

(١) تقصد برمجية هذا المعيار [المترجم].

المتجانس للنتائج عندما لا يكون هناك قدر كاف من التوافق الوظيفي بين نظم البحث المختلفة، فقد يتوافر في أحد النظم - على سبيل المثال - كشاف " المؤلفين author " (ويستخدم الخاصية ١٠٠٣) ، بينما يتوافر لنظام آخر هذا النوع من الكشافات ولكن باسم كشاف الأسماء name" (ويستخدم الصفة ١٠٠٢)، فعندما يرسل النظام الأول استفساراً معتمداً على بروتوكول Z٣٩,٥٠ عن "المؤلف" ، فستتوقف نتيجة البحث على مدى تعرف النظام الثاني على طبيعة السؤال ، ومن ثم فإذا رفض البحث [من قبل هذا النظام الثاني] لعدم وجود كشاف مواز، فإن الباحث بالتالي يكون قد فقد مواد ذات صلة بموضوع بحثه من ناحية ، وإذا قام النظام الثاني بربط البحث بكشاف "الأسماء " فسوف يحصل الباحث على مواد مسترجعة خاطئة ، أي ليس لها علاقة بموضوع بحثه من ناحية أخرى.

وهناك اتجاه واعد يهدف إلى الارتقاء بعمليات البحث البيني يتمثل في تبني ما يعرف بالملف التفصيلي للسمات أو بروفايل زد ٣٩,٥٠ profile Z٣٩,٥٠ ،حيث يحدد هذا الملف وظائف معينة ينبغي أن تقدمها أو تدعمها برمجية عميل المعيار Z٣٩,٥٠ وخادمه، ويحدد كذلك تعريفات أو مواصفات معينة ينبغي عليها استخدامها ، هذا فضلاً عن ضرورة توافر كشافات محددة لإجراء عملية البحث، بحيث إن أي مجموعة من الحاسبات الخادمة للمعلومات information servers تتوافق وظائفها مع الوظائف التي يحددها البروفايل تكون قادرة على إتاحة خدمات بحث بيني جيدة ومعقولة. ومن البروفايلات المسجلة دولياً في هذا الصدد

، بروفايل الباث Bath profile " الخاص بالتطبيقات المكتبية[1]، كما تقوم المنظمة الوطنية لمعايير المعلومات NISO بتطوير بروفايل وطني أمريكي معتمد على بروفايل الباث .

ومع أن تطوير البروفايل قد تركز حتى الآن على البيانات الببليوجرافية ، فإن ما أثير من قضايا قد لا تختلف عما قد يثار حول الأشكال الأخرى لما وراء البيانات، وقد رعى اتحاد المكتبات الرقمية مشروعاً تجريبياً لعمل خدمة بحث معتمدة على بروتوكول Z٣٩,٥٠ لمجموعات موزعة من تسجيلات الوصف الأرشيفي المرمز EAD، واتفق المشاركون في المشروع على دعمهم لإيجاد مجموعة موحدة من الكشافات أو " نقاط الوصول المشتركة common Access Points " للبحث عبر الأنظمة المختلفة ، هذا بجانب الكشافات المتاحة للاستخدامات المحلية. وبصرف النظر عما أحاط بمعدلات الاسترجاع المتناغم أو المطرد من قيود نتيجة للممارسات المتباينة من قبل الجهات المختلفة لتطبيق خطط ترميز الوصف الأرشيفي المرمز EAD mark up فقد ثبت للمشروع أن نموذج البحث الموزع كان مجدياً[2].

وتحيط ببروتوكول Z٣٩,٥٠ قيود أخرى يصعب تجاوزها عن طريق

---

(١) The Bath Profile : An International Z٣٩,٥٠ Specification for Library Applications

Resource Discovery, Release ١,١, available at

http://www.ukoln.ac.uk/interopfocus/bath/l.I/. Accessed ١١ June ٢٠٠٢

(٢) Mackenzie Smith, "DFAS: The Distributed Finding Aid Searh System," D-Lib Magazine ٦, no. l (January ٢٠٠٠), available at

.http://www.dIib.org/dlib/januaryOO/OIsmith.html.Accessed ١١ June ٢٠٠٢

البروفايل ، منها أن واجهات هذا البروتوكول نادراً ما تدعم كل الإمكانات الوظيفية التي توفرها الواجهات الأصلية للنظم المختلفة ، بل إن إمكانات هذه الواجهات قد لا تتناسب والبحوث المعقدة ، كما أن من المعروف على نطاق واسع أن البروتوكول Z٣٩,٥٠ لا يحقق معدلات نجاح عالية في حالة البحث الموسع لعدد كبير جداً من المصادر. ومن ثم ضرورة تأسيس اتصالات قوية بين الحاسب العميل وجميع الحاسبات الخادمة ، فقد تتأخر عملية التعرف أو الاتصال وخاصة عند عدم توافر حاسب واحد أو أكثر من الحاسبات الخادمة مما يضطر الحاسب العميل إلى الانتظار حتى تنتهي المهلة التي يجب أن يُرى خلالها الحاسب الخادم.

وهناك مشكلة أخرى تتصل بهذا البروتوكول تتمثل ببساطة في محدودية عدد مصادر المعلومات التي يمكن إتاحة الوصول إليها عن طريقه ، ذلك لأن [البرمجيات] الخادمة للبروتوكول (Z٣٩,٥٠ servers) لا يسهل تطويرها وصيانتها بسبب تعدد الوظائف التي يدعمها البروتوكول من جانب ، وكذلك بسبب العمر الزمني لهذا المعيار من جانب آخر ؛ ففي سياق هذا البروتوكول يتطلب الأمر أن ترسل البيانات - على سبيل المثال - في خطة ترميز معقدة لم تعد متداولة الآن على نطاق واسع، بل لم تعد تعرفها سوى قلة من البرامج. لهذه الأسباب وغيرها يحجم كثير من مزودي خدمات المعلومات المباشرة عن إتاحة خدماتهم اعتماداً على هذا البروتوكول.

وقد شكلت جماعة فرعية من مجموعة مطبقي بروتوكول Z٣٩,٥٠ implementors Group(ZIG) للتعامل مع بعض هذه الإشكالات ، وانتهت إلى استحداث جيل جديد من هذا البروتوكول أطلقت عليه" الجيل الدولي

التالـي من بروتوكول زد٣٩,٥٠ "Z٣٩,٥٠-International Next Generation" ، أو مـا يعـرف اختـصاراً بــ" زنـج (ZING)"[١] . وقـد حـاول هـذا الـتحديث استثمار إمكانـات أكثر التقنيات تآلفاً مع الـوب، مثل: بروتوكول نقل الملف الفائق HTTP ، و لغة الترميز الموسعة XML ، وبروتوكول سواب SOAP (بـروتوكول بـسيط لتـبادل الرسـائل مبنـي علـى لغـة الترميـز الموسـعة) ، والاستغناء عن أقل إمكانـات المعيار استخداماً وأكثرها تقادماً. وثمة بارقة أمل أن تلك التحديثات لآليات هذا البروتوكول، مع الاحتفاظ بما تتميز به دلالاتـه اللغوية المجردة من قدرات هائلة ، تدفع مقدمي خدمات المعلومات لتبني هذا البروتوكول .

وباختصار ، يمكن القول إن المزية الرئيسية للبروتوكول Z٣٩,٥٠ تتمثل فـي إعفائـه بـرمجية الحاسـب العمـيل مـن الحاجـة إلـى معـرفة بـروتوكول الوصـول ،ومن الحاجـة إلـى معرفة صيغة البحث لكل نظام بعيد يطمح أن يجري بحثاً فيه ، أما نقطـة الـضعف الكبـرى فتتمثل في أن الحاسب العميل لبــروتوكول Z٣٩,٥٠ لا يـتمكن إلا مــن الـتخاطب مــع [بــرمجية] الخــادم للبــروتوكول Z٣٩,٥٠ ، ومــن المعــروف أن بعــض خــدمات المعلــومات المباشـرة لا توفر [برمجيات] خادمة للبروتوكول Z٣٩,٥٠، وقد طورت عدة بـرمجيات للـتعامل مـع هـذه المـشكلة الأخيـرة عـن طـريق إلحـاق اتـصالات بروتوكول Z٣٩,٥٠ طرقًا أخرى لبحث خدمات المعلومات.

---

(١) "ZING Z٣٩,٥٠-International: Next Generation": available at http://www.loc.gov/z٣٩٥٠/agency/zing/zing.html. Accessed ١١ June ٢٠٠٢

ويعتمد العديد من خدمات المعلومات المتاحة على الوب على استخدام محددات بحث تمرر في الجزء الخاص بالاستفسار من المحدد الموحد لموقع المصدر URL، فمثلاً عند البحث عن المعلومات المتصلة بموضوع " حمى القش hay fever "، ستأخذ الكلمة الأساسية العامة التي تعبر عن هذا الموضوع - أي حمى القش - صيغاً مختلفة عند البحث في الخدمات التالية : بب ميد PubMed، وجوجل Google ، و فهرس مكتبة جامعة شيكاغو the University of Chicago Library catalog - وذلك على النحو التالي :

- عند البحث في خدمة بب ميد ستأخذ الصيغة التالية:

http://www.ncbi.nlm.gov/entrez/query.fcgi?cmd=search&db=pubMed&term=hay+fever

- وعند البحث في جوجل ستأخذ الصيغة التالية:

http://www.google.com/search?q=hay+fever

- وعند البحث في فهرس مكتبة جامعة شيكاغو ستأخذ الصيغة التالية:

http://ipac.lib.uchicago.edu/ipac/ipac?uofc=on&db=uofc&sp١=.gk&se=hay+fever&tm=summary

على أن تقوم البرمجيات المصممة لأغراض البحث الموسع بتلقي تساؤل البحث الذي يقدمه المستفيد ، ثم تترجمه إلى أي من استراتيجيات البحث الثلاثة ، ومن ثم إرسال كل استراتيجية إلى الخدمة المستهدفة المناسبة ، وعادة ما يكون لمثل هذه البرمجيات نصوص أو خطوط scripts

مستقلة أو نظم فرعية وظيفية آلية خاصة بكل خدمة مستهدفة تستطيع البحث فيها . وبالطبع فإن الاستجابة الطبيعية من قبل خدمة المعلومات المتاحة على الوب لاستراتيجيات البحث القائمة على المحدد الموحد لموقع المصدر URL ستكون في شكل قائمة مرتبة بالنتائج المطابقة لاستراتيجية البحث، وليست في شكل استدعاء أو استرجاع لملف التسجيلات الأصلية، ولذلك فإن هذه البرامج عادة ما يجب عليها القيام بعملية "استقصاء أو مسح screen scraping "(١) أو التعرف إلى شكل المواد المسترجعة من أجل فصل النتائج المطابقة الصحيحة المسترجعة عن المواد غير المطابقة، إلا أن مثل تلك البرامج غالباً ما تكون متأثرة بالتغيرات البسيطة التي تحدث في صياغة المخرجات ،وكذلك بالتغييرات الكبيرة التي تحدث في صيغة البحث الخاصة بأي خدمة معلومات معينة، ونتيجة لذلك فإنها تحتاج إلى مستوى عال من الصيانة المستمرة أكبر من الوضع الذي تكون عليه الخدمات المعتمدة على بروتوكول Z٣٩,٥٠.

وهنالك عدد متزايد من خدمات المعلومات التي تتيح بوابات لغة الترميز الموسعة XML gateways ، وهي عبارة عن واجهات تقبل

---

(١) يستخدم هذا المصطلح ليدل على قيام أحد البرامج الآلية باستخراج أو اشتقاق البيانات التي يعرضها برنامج آخر ، ويسمى البرنامج الذي يقوم بعملية الاشتقاق هذه ماسح الشاشة screen scraping ، وهناك عدة استخدامات لهذا المصطلح أهمها : data scraping , data extraction, web scraping. ,page scraping , HTML scraping ، [المترجم عن موسوعة Wikipedia].

استراتيجيات البحث المصاغة وفقاً لهذه اللغة XML-formatted queries ، وتسترجع النتائج في شكل تسجيلات مهيأة وفقاً لهذه اللغة كذلك XML-formatted records. ومع أن الصيغ الإجبارية للاستفسارات، كما هو الحال مع مؤشرات البحث القائم على المحدد الموحد لموقع المصدر URL-based search parameters ، لا تزال تتفاوت من نظام لآخر، فإن هذه الأساليب تتمتع بمزية إرسال النتائج في شكل تسجيلات، ولذلك فلا تكون هناك حاجة لعملية الاستقصاء.

وتسعى كثير من تطبيقات البحث الموسع إلى الجمع بين إمكانية إجراء بحوث معتمدة على بروتوكول Z٣٩,٥٠ عن طريق نظم فرعية وظيفية مخصصة لتنفيذ عمليات البحث القائم على المحدد الموحد لموقع المصدر URL-based searching ، وبحث بوابات لغة الترميز الموسعة ،بالإضافة إلى بعض أساليب البحث الأخرى من أجل توفير أكبر عدد ممكن من الخدمات المستهدفة والتي تشمل منتجات موردي نظم المكتبات ، مثل : Ex Libris's Metalib ، و Endeavor's Encompass ، بالإضافة إلى مشروعات التطوير المحلية ، مثل : أداة البحث الخاصة بمكتبة كاليفورنيا الرقمية والمعروفة بـ" سرش لايت " the California Digital Library's Searchlight (http://searchlight.cdlib.org/cgi-bin/searchlight) ، التي تستخدم كلاً من البروتوكول Z٣٩,٥٠ ، وأساليب بحث الوب لتوفير خدمة بحث موسع في فهارس المكتبات وقواعد بيانات الدوريات وغيرها من مصادر المعلومات الأخرى.

نقاط المقاربة أو التحويل البيني[1] **Crosswalks** :

تستخدم نقاط المقابلة أو التحويل البيني لتيسير عمليات التشغيل البيني أو التوفيق بين خطط ما وراء البيانات المختلفة ، وتعد هذه النقاط عملية مقاربة أو مقابلة mapping بين عناصر ما وراء بيانات في إحدى الخطط بعناصر خطة أخرى ، حيث تحدد خطة المقابلة التي تتبناها مكتبة الكونجرس - مثلاً - بين دبلن كور ومارك أن عنصر " المساهم contributor" في خطة دبلن كور غير المقيد يقابل الحقل " ٧٢٠ ذا المؤشرات الخالية" في مارك ، وهذا يعني أن القيمة التي تعطى لعنصر" المساهم" يجب أن تظهر في الحقل الفرعي " a " للحقل ٧٢٠ ، كما أن كلمة "مساهم contributor " يجب أن تظهر بهذا الصيغة الحرفية في الحقل الفرعي" e " للحقل ٧٢٠[2]. والشكل رقم (١/٤) هو فقرة طبق الأصل من خطة مكتبة الكونجرس للمقابلة بين كل من دبلن كور ومارك وخدمة محدد موقع المعلومات الجغرافية GILS فيما يتصل بعنصر "المساهم".

ويمكن القول إن نقاط التحويل البيني هي روابط خطية lateral (أي في اتجاه واحد) تربط خطة معينة بخطة أخرى، ولذلك فإن ربط الخطة (أ) بالخطة (ب) من ناحية ، والخطة (ب) بالخطة (أ) من ناحية أخرى ، يتطلب وجود نقاط

---

(١) المعنى اللغوي لهذا الكلمة هو الممرات أو ممرات عبور المشاة ، أو الخطوط التي تحدد تقاطع الطرق . وقد رأيت التحرر من هذا المعنى والإفادة من مدلوله لاختيار كلمة " نقاط التحويل البيني أو نقاط المقابلة [المترجم].

(١) Library of Congress Network Development and MARC Standards Office, "Dublin Core/MARC/GILS Crosswalk": ٢٠٠١, available at http://lcweb.loc.gov/marcldccross.html. Accessed ١٢ June ٢٠٠٢.

تحويل بيني مستقلة . وبشكل عام ، وحتى مع وجود نقطتي تحويل بيني فإن عملية الربط مكتملة الدورة من غير المتوقع أن تتم دون فقدان قدر ما من المعلومات أو اضطراب بعض المعلومات ، بمعنى أنه إذا أجري ربط أو مقابلة لتسجيلة ما وراء بيانات من الخطة (أ) إلى الخطة (ب) ، ثم مرة أخرى إلى الخطة (أ) ، فمن غير المحتمل أن تأتي التسجيلة الناتجة عن عملية المقابلة هذه مماثلة تماماً للتسجيلة الأصلية.

وفي هذا الصدد طورت نقاط تحويل بين خطط ما وراء البيانات الرئيسة استجابة لأغراض وصف مصادر المعلومات، حيث ترعى مكتبة الكونجرس الآن تطوير نقاط تحويل بين مارك ٢١ وبعض الخطط الأخرى ، مثل خطة التحويل بينه وبين دبلن كور ، ومعايير المحتوى الخاصة باللجنة الفيدرالية للبيانات الجغرافية Federal Geographical Data Committee (FGDC) لما وراء البيانات المكانية Geospatial ، وخدمة محدد مواقع المعلومات الكونية GILS،

Creator
An entity primarily responsible for making the content of the resource.

MARC 21:

Unqualified:

- 720 ##$a (Added Entry--Uncontrolled Name/Name) with $e=author

Qualified:

- Personal: 700 1#$a (Added Entry--Personal Name) with $e=author
- Corporate: 710 2#$a (Added Entry--Corporate Name) with $e=author
- Conference: 711 2#$a (Added Entry--Conference Name) with $e=author
- Role: 720 ##$e (Added Entry--Uncontrolled Name/Relator term
- Role (Personal): 700 1#$e (Added Entry--Personal Name/Relator term)
- Role (Corporate): 710 2$e (Added Entry--Corporate Name/Relator term)

Note: The above qualifiers have not been approved by DCMI.

GILS:

- Originator

الشكل رقم (٤/١)

جزء مقتبس من نقطة تحويل مكتبة الكونجرس بين دبلن كور مارك وجيلس

وخطة تبادل المعلومات المباشرة المعروفة بأونيكس[١] ONIX. وهناك نقاط

(١) خطة طورتها مجموعة من الناشرين لتبادل المعلومات التجارية في صيغة إلكترونية [المترجم].

تحويل بيني إلى خطط ما وراء بيانات أخرى ومنها تضطلع بها الجهات المسؤولة عن تلك الخطط. كما أن هناك بعض الخطط الأخرى مثل خطة الفئات الأساسية لجمعية المصادر المرئية VRA Core ، وما وراء بيانات الكيان التعليمي لجمعية المهندسين الكهربائيين ومهندسي الإلكترونيات IEEE ، تتضمن كجزء من توثيقهما نقاط تحويل من عناصرها إلى الخطط الأخرى.

وتتمثل الاستخدامات الرئيسة لنقاط التحويل في أنها تعد بمثابة مواصفات أساسية تقوم عليها عملية التحويل الفعلي للتسجيلات من خطة ما وراء بيانات إلى خطة أخرى بغرض تبادل هذه التسجيلات، أو المشاركة في الفهارس الموحدة، أو تجميع ما وراء البيانات. كما يمكن أيضاً لمحركات البحث أن تستخدم خطط التحويل البيني لبحث حقول تشتمل على المحتوى نفسه أو محتوى مشابه في قواعد بيانات مختلفة، وهنالك جانب آخر من جوانب استخدام نقاط التحويل لم يقدر قدرها برغم أهميته، وهو مساعدة الإنسان على استيعاب خطط ما وراء البيانات الجديدة، إذ يمكن لشخص ما فرض عليه أن يتعامل مع خطة غير مألوفة له أن يستخدم نقطة التحويل المعدة للربط بين تلك الخطط التي لم يألف التعامل معها وخطة أخرى يألف التعامل معها بشكل أكبر في استنباطاته لدلالات عناصر ما وراء البيانات واستخداماتها.

ومع ذلك تحيط بمسيرة تطوير نقاط التحويل البيني عدة عقبات نتيجة لغياب الصيغة المعيارية التي تحكم تمثيل خطط ما وراء البيانات ، والتي نتج عن غيابها ذلك التفاوت الواضح فيما بين الخطط سواء في تحديد

خصائص عناصر بياناتها ،أو في مسميات هذه العناصر[١]؛ ولذلك فإن أولى مراحل إنشاء خطط التحويل البيني ينبغي أن تركز على وضع كلتا الخطتين: المصدرية والمستهدفة ، في صيغ متشابهة ، بحيث يعبر عن الخصائص المتماثلة بطريقة مشابهة ، ثم تأتي بعد ذلك مرحلة التوفيق بين عناصر البيانات المتفاوتة نفسها، فقد يكون لإحدى الخطط - على سبيل المثال - عنصر موضوعي متكرر ، بينما يكون لخطة أخرى حقل موضوعي غير متكرر ، وحسب المصطلحات الموضوعية المتعددة التي تُضمّن داخل هذا الحقل أن يفصل بينها بعلامة الفاصلة المنقوطة[٢]، من هنا كان على خطط التحويل البيني أن تبين الكيفية التي تعامل بها مثل هذه المقابلات التي قد تكون من عنصر واحد في خطة واحدة إلى العنصر المقابل في عدة خطط ، أو من عنصر في عدة خطط إلى عنصر واحد في خطة واحدة؛ فقد لا يكون لأحد العناصر في خطة مصدرية معينة عنصر مقابل في خطة أخرى مستهدفة ، أو أن هذا العنصر الواحد يقابله عنصران أو أكثر من العناصر المستهدفة ، كما أنه قد تكون هناك اختلافات في تمثيل البيانات (كما هو الحال بالنسبة للأسماء المقلوبة)، أو في قواعد المحتوى (كتلك الاختلافات حول مدى الالتزام باللغات المقيدة).

أما أكثر القضايا التي يصعب حلها فهي تلك التي تتعلق بالاختلافات الدلالية الحقيقية ؛ ففي بداية استحداث دبلن كور - مثلاً - لم يكن هناك حقل

(١) Margaret St. Pierre and William P. laplant, Jr., "Issues in Crosswalking Content Metadata Standards:" ١٩٩٨, available at http://www.niso.org/press/whitepapers/crsswalk.html.Accessed ١٢ June ٢٠٠٢.

(٢) تعرف في مجال الفهرسة بشبه الشارحة [المترجم].

في صيغة مارك يمكن أن يقابل أو يضاهى به عنصر المنشئ creator في دبلن كور بشكل دقيق ، ذلك أن الحقول الخاصة بالأسماء في مارك تعرف الأسماء في حالتين، إما مداخل رئيسة وإما مداخل إضافية ، وهما من المفاهيم المعروفة في القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية ، لكنهما غير معروفين بأي حال من الأحوال في دلالات دبلن كور. ولكي يمكن تطوير خطط تحويل من دبلن كور إلى مارك استلزم الأمر إضافة حقل جديد في صيغة مارك الأمريكي USMARC ، هو الحقل "٧٢٠ " (الخاص بالمدخل الإضافي - اسم غير مقيد)(١).

مسجلات أو مكاتب تسجيل ما وراء البيانات **Metadata Registries** :

مكتب تسجيل ما وراء البيانات هو أداة لتسجيل معلومات موثقة عن عناصر ما وراء البيانات من مصادر متعددة؛ فبتسجيل أسماء عناصر ما وراء البيانات وتعريفاتها وخصائصها ، تيسر مكاتب التسجيل التعريف بعناصر ما وراء البيانات وإتاحة الفرصة لإعادة استخدامها ، والتوفيق بينها، وفي ظل التزايد المتنامي لاستخدام خطط ما وراء البيانات في مختلف المجالات العلمية ، يتزايد الاهتمام بمكاتب تسجيل ما وراء البيانات على اعتبارها أدوات لإدارة البيانات.

ومن الملاحظ أنه لا يوجد من مسجلات ما وراء البيانات العملية إلا القليل جداً ، كما أن معظم المسجلات العاملة الآن أو ما هو قيد التطوير منها

---

(١) [MARBI], Proposal No. ٩٦-٢ (١٩٩٦), available at http://lcweb.loc.gov/marc/marbi/١٩٩٦/٩٦-٠٢.html. Accessd ١٢ Hune ٢٠٠٢.

يعتمد على معيار اللجنة الدولية للتقنيات الإلكترونية التابعة للأيزو - المعيار رقم ١١١٧٩ ISO/IEC - الخاص بتوصيف وتقييس عناصر البيانات ، حيث تغطي الأجزاء المختلفة من هذا المعيار السمات الرئيسة لعناصر البيانات ، كما تصف كيفية صياغة أسماء عناصر البيانات وتعريفاتها، وكيفية إنشاء هيئات التسجيل ومكاتب تسجيل ما وراء بيانات. ومن الملاحظ أن تطوير هذا المعيار لم يأت تلبية لكثرة خطط ما وراء البيانات على النحو الذي أشرنا إليه ، بل جاء بدافع إدارة نظم قواعد البيانات المتعددة المتواجدة بداخل جهة معينة ، أو المتواجدة في داخل مجموعة من الجهات والتي يكون لكل منها خطتها الخاصة بها ، أو يكون لها مجموعة من العناصر المحددة . وقصد من مكاتب تسجيل ما وراء البيانات أن تسمح لبعض الجهات عالية المستوى بتجميع كل تلك المعلومات بطريقة متجانسة وبما يسمح بتعريف عناصر البيانات وفهمها وإعادة استخدامها.

وتعد قاعدة المعرفة للمعهد الأسترالي للصحة والرفاهية the Australian Institute of health and welfare knowledgebase التي تضم تعريفات للعناصر ذات الصلة بمجالات الصحة وخدمات المجتمع وإعانات الإسكان (http://www.aihw.gov.au/knowledgebase/) ، واحدة من أشهر مكاتب تسجيل ما وراء البيانات المبنية على المعيار ١١١٧٩ ISO/IEC ، كما يعد مكتب تسجيل البيانات البيئية لوكالة حماية البيئة the environmental Protection Agency's Environmental Data Registry أو ما يعرف اختصاراً بـ إي دي آر EDR (http://www.epa.gov/edr/) ، أكبر مطبقي هذا المعيار ١١١٧٩ ISO/IEC في الولايات المتحدة الأمريكية ؛ حيث يقوم هذا

المكتب بوصف البيانات البيئية الموجودة داخل حدود الوكالة وخارجها ، كما يشتمل (حسب إحصاء سبتمبر ٢٠٠٢م) على ٩٧٥١ عنصر بيانات ، جمعت من ١٤١٩مصدرًا للمعلومات، قدمتها خمس وخمسون جهة مختلفة. ومن ثم فعندما يتم البحث عن المقطع "zip" - على سبيل المثال - يسترجع أكثر من مئة عنصر من عناصر ما وراء بيانات المعرفة في عدة قواعد بيانات مختلفة، ومن ضمن ما يسترجع عنصر البيانات المسمى" ZIP_CODE " (المعرف ١:٢٤٥٥٦-١) الذي يشير إلى رمز "خطة تطوير المنطقة the Zone Improvement Plan " ، مع الامتداد المكون من الخانات الأربعة الممثلة لعنوان موقع المؤسسة في الواقع الفعلي ، والمأخوذ من جدول " DNB Company" في قاعدة بيانات Dun and Bradstreet التابعة لقاعدة المعلومات البيئية المعروفة بإنفايروفاكتس Envirofacts ، كذلك يسترجع العنصر " ZIP_CODE" (المعرف ١:٢٤٥٥٦-١) الذي يشير إلى رمز " خطة تطوير المنطقة " في العنوان المرتبط بخدمة العنوان البريدي " المأخوذ من جدول " RCR Mailing Location " في نظام معلومات صيانة المصادر واستعادتها في قاعدة المعلومات البيئية إنفايروفاكتس Envirofacts نفسها .

ويمكن القول بشكل أكثر دقة إن مكاتب التسجيل التي طورت لخطط ما وراء البيانات بغرض وصف مصادر المعلومات لم تأت ملتزمة تماماً بالمعيار ISO/IEC ١١١٧٩ بقدر ما جاءت مراعية له. وثمة ثلاثة اتجاهات مختلفة لمكاتب التسجيل تبنتها المشروعات الثلاثة التالية: مشروع رودز ROADS ، ومشروع مبادرة دبلن كور لما وراء البيانات DCMI ،

ومشروع ديزاير DEISRE.

ومشروع رودز ROADS (أو تنظيم المصادر واستكشافها في الخدمات القائمة على أساس الموضوع Resource Organization And Discovery in) Subject-based Services) هو مشروع لبرنامج المكتبات الإلكترونية للجنة المشتركة لنظم المعلومات Joint Informatiom Systems Committee (JISC) في المملكة المتحدة، ويسعى هذا المشروع إلى توفير برمجيات تسمح للمشاركين بتجهيز البوابات الموضوعية القائمة على تسجيلات ما وراء البيانات المدخلة وفقاً لقوالب معيارية جاهزة ، يتوافر منها قوالب خاصة بالوثائق والصور والمواد السمعية والتجميعات وغيرها من أنواع المواد الأخرى. ومكتب تسجيل ما وراء البيانات في هذا المشروع (رودز) هو ببساطة عبارة عن قائمة بالقوالب الجاهزة مع العناصر التي تشتمل عليها (http://www.ukoln.ac.uk/metadata/roads/templates). وعلى ذلك فإن هذا المكتب لا يعد بالمفهوم الدقيق للمعيار ISO/IEC١١١٧٩ مسجلاً ؛ ذلك لأنه لا يوفر إمكانات بحث باسم العنصر، كما أنه لا يقدم إمكانية مقارنة استخدام العنصر في عدة قوالب مختلفة، ولكنه مع ذلك يوفر بالفعل مكاناً تسجل فيه الإصدارات الرسمية لمختلف القوالب الجاهزة المختلفة.

أما تطوير مكتب تسجيل ما وراء البيانات لدعم مبادرة دبلن كور لما وراء بيانات فيعد مشروعاً نشطاً تبنته مجموعة عمل مكتب تسجيل دبلن كور DCMI (http://www.dublincore.org/groups/registry). وقد عنيت هذه المجموعة بتطوير أداتين: الأولى خاصة بنظام إدارة اللغات المقيدة ، والأخرى خاصة بمكتب التسجيل. ويتمثل الهدف من وراء تطوير الأداة الأول في

المساعدة في إدارة عناصر دبلن كور وتطويرها من خلال تقديم معلومات كافية عن جميع المصطلحات الحديثة والقديمة وتوفير تعريفات لهذه المصطلحات ، وتوفير أساليب إقرار المصطلحات الجديدة ، هذا فضلاً عن بعض الوظائف الأخرى المشابهة. أما الأداة الثانية الأخرى - أي مكتب التسجيل - فهي أداة للمستخدم النهائي (سواء أكان الإنسان أم البرمجية) يمكنه بموجبها الحصول على معلومات وافية عن مصطلحات دبلن كور، وعن استخدامها ، وعن العلاقات القائمة بينها، وفضلاً عن ذلك يوفر مكتب التسجيل واجهة متعددة اللغات ، وتوصيفاً للمصطلحات بلغات متعددة.

أما مكتب تسجيل ما وراء بيانات لمشروع ديزاير (http://desire.ukoln.ac.uk/registry) فقد تبنى اتجاهاً أوسع، ومشروع ديزاير DESIRE (أو تطوير خدمة أوربية لمعلومات البحث والتعليم Development of a European Service for Information on Research and Education) مشروع تعاوني بين عدة جهات من هولندا والنرويج والسويد والمملكة المتحدة، جرى تمويله خلال الفترة من عام ١٩٩٨ حتى ٢٠٠٠م . وعلى عكس الغالبية العظمى من مكاتب التسجيل القائمة على المعيار ISO/IEC ١١١٧٩ ، فإن مكتب تسجيل ديزاير قد صمم بهدف إدارة عناصر ما وراء البيانات المأخوذة من أسماء مميزة namespaces (أو خطط) متعددة . وتحقيقاً لهذا الهدف فقد قام ببناء نظام لعرض ما وراء البيانات demonstrator مشتملاً على العناصر المأخوذة من ثلاث إصدارات مختلفة لدبلن كور ، ومن إحدى الصيغ الموسعة من دبلن كور التي استخدمها

مشروع بيبلنك BIBLINK [1] ، ومن قوالب جاهزة مختارة من مشروع رودز، هذا فضلاً عن عناصر مأخوذة من بعض الخطط الأخرى.

ولعل من أهم الجوانب المثيرة في هذا الاتجاه الذي يتخذه مشروع ديزاير يتمثل في اختباره لإمكانية توليد نقاط التحويل البيني آلياً عن طريق استخدام مسجل الدلالات الأساسية للأيزو (ISO Basic Semantics Register(BSR)، وهو عبارة عن توليفة معترف بها دولياً من عناصر بيانات مصممة من أجل تطوير نظم في بيئة متعددة اللغات. ففي مشروع ديزاير وبدلاً من محاولة عمل مقابلات أحادية الطرف unilateral mappings من جميع عناصر إحدى الخطط إلى جميع عناصر أي خطة أخرى ، يقابل كل عنصر بمفهوم دلالي محايد a neutral semantic concept (يسمى بلغة الأيزو بالوحدة الدلالية semantic unit)، بحيث يمكن الربط أو المقابلة بين أي خطتين - ولتكن الخطة (أ) والخطة (ب) - عن طريق مقابلة عناصر الخطة (أ) بالوحدات الدلالية المقابلة في مسجل الدلالات الأساسية للأيزو، ثم مقابلة الوحدات الدلالية في مسجل الدلالات الأساسية بالعناصر المقابلة لها في الخطة (ب). ومن الواضح أن هذا المفهوم يبدو مألوفاً لمن لديه معرفة بالبروتوكول Z٣٩,٥٠.

وبقدر ما هنالك من اهتمام باستحداث مسجلات مفهومة آلياً machine-understandable registries ، هنالك اتجاه بحثي نشط لاستحداث مسجلات

(١) بب لينك BIBLINK : مشروع مول من الاتحاد الأوربي لتقوية العلاقة بين الهيئات الببليوجرافية الوطنية وناشري المواد الإلكترونية [المترجم].

مفهومة بشرياً human-understandable (انظر الجزء الخاص بالوب الدلالية في الفصل الخامس). ومع أن مسجلات ما وراء البيانات لا تزال في بداياتها ، فمن المتوقع أن تنضج تقنياتها ، كما ستزداد أهميتها مع مضي الوقت استجابة للمشكلات المتفاقمة التي تثيرها خطط ما وراء البيانات المتعددة بكل إصداراتها وتنوعاتها.

معوقات التشغيل البيني :

بصرف النظر عن الطريقة المتداولة في البحث عبر المصادر المتعددة ، فإن أوجه التفاوت الأساسية المحيطة بما وراء البيانات سوف تثير صعوبات في الاسترجاع والعرض ، بل يمكن القول بأنه كلما ازدادت درجة التفاوت أو عدم التجانس بين خطط ما وراء البيانات ، تفاقمت مشكلات الاسترجاع. وثمة عدة قضايا تثار أكثر من غيرها في هذا السياق ، ومن أهم ما تجدر الإشارة إليه من هذه المشكلات ما يلي :

الفروق الدلالية : فليس هنالك بالضرورة تطابق كامل بين معاني عناصر ما وراء البيانات من الخطط المختلفة، بل قد تكون الفروق كبيرة بين هذه الخطط (كما هو الحال عندما لا يكون لأحد العناصر في خطة معينة عنصر مقابل في الخطط الأخرى) ، وقد تكون هذه الفروق طفيفة ؛ فعنصر العنوان - على سبيل المثال - في دبلن كور عبارة عن اسم يعطى لمصدر المعلومات ، بينما عنصر العنوان نفسه Title Proper (الحقل ٢٤٥) في كل من القواعد الأنجلو أمريكية ومارك AACR٢/MARC ، لا يمكن تعيينه إلا باتباع مجموعة قواعد مقننة ومحددة تحديداً جيداً، ومن ثم فإن مدى تكافؤ عنصري العنوان في هاتين الخطتين يعتمد على كم الأخطاء التي يستطيع

المرء أن يقبلها. وفي نموذج آخر نجد أن خدمة محدد مواقع المعلومات الكونية GILS تشتمل على عنصر " المؤلف Author " ، بينما تشتمل كل من دبلن كور والمجموعة الأساسية لجمعية المصادر المرئية VRA Core على عنصر أكثر عمومية ، هو عنصر "المنشئ creator "، أما القواعد الأنجلو أمريكية ومارك AARC٢/MARC فتشتملان على عنصري "المداخل الرئيسية والمداخل الإضافية "، وكلاهما يعد أكثر عمومية من ناحية (لأنهما يستوعبان أنواعاً مختلفة من الأدوار)، وأكثر تخصيصاً من ناحية أخرى (من حيث وجود قيود على هوية الشخص أو الجهة التي ستظهر كمدخل رئيسي أو ستظهر كمدخل إضافي)، كما أن الوصف الأرشيفي المرمز EAD يشتمل على عنصر " المؤلف "، ولكنه يستخدم لوصف الشخص القائم بالوصف الأرشيفي نفسه، وليس الشخص الذي يعزا إليه تكوين مجموعة المقتنيات التي يتم وصفها عن طريق الوصف الأرشيفي المرمز.

تفاوت التطبيق : إذ تختلف طريقة الوصف باختلاف المجتمع الذي يقوم بالوصف ، فما وراء البيانات التي ينشئها الأفراد أو المؤسسات كالمكتبات والأرشيفات والمتاحف والجمعيات التاريخية تختلف بعضها عن بعض حتى في العناصر الأساسية؛ فأمناء المكتبات - على سبيل المثال - يصرون بشكل عام على كتابة عنصر " العنوان "، وقد يضطرون إلى صياغة عناوين للأعمال التي ليس لها عناوين، أما أمناء المجموعات المتحفية museum curators فيفضلون - في المقابل - حذف عناوين الأعمال الفنية ثلاثية الأبعاد، ويعتمدون بدلاً عن ذلك على وصف الموضوع ، كما يلاحظ أن مبادئ وقواعد الوصف الأرشيفي تختلف عن قواعد الوصف

الببليوجرافي ، كما أنه لا يوجد في معينات البحث الأرشيفية مفهوم " المؤلف " ، بالرغم من إمكانية عمل تيجان للأسماء الشخصية وأسماء الهيئات وأسماء العائلات حيثما تكون ثمة علاقة بينها(١).

اختلاف طريقة التمثيل: فحتى عندما تكون تعريفات العناصر متماثلة ، يمكن تسجيل البيانات بأشكال مختلفة حسب قواعد التمثيل ؛ فمثلاً إذا كانت إحدى مجموعات تسجيلات ما وراء البيانات تعرض أو تمثل المؤلف في الصيغة التالية " Public, Johan Q. " ، وتعرضه مجموعة أخرى بالصيغة " Public, J.Q."، فإن البحث بالكلمتين الأساسيتين " Johan Public "، ينتج عنه استرجاع التسجيلات التي وردت فيها الصيغة الأولى دون الصيغة الثانية . في هذا الصدد تسعى بعض واجهات البحث الذكية إلى تخفيف حدة مشكلات الاختلافات العامة ؛ فمثلاً هناك واجهة بحث موسع تسمى فلاش بوينت " Flashpoint " طورتها مختبرات لوس آلاموس الوطنية Alamos National Laboratories ، تنبه المستفيد إلى إعادة البحث عن المؤلف باستخدام الأحرف الاستهلالية من اسمه في حالة عدم استرجاع أي نتائج عندما يبحث عنه بصيغة هجاء اسمه الأول كاملاً(٢).

---

(١) William Garrison, “Retrieval Issues for the Colorado Digitization Project’s Heritage Database,” D-Lib Magazine ٧, no. ١٠ (October ٢٠٠١), available at

http://www.dIib.org/dlib/octoberOi/garrison/iOgarrison.html. Accessed ١٢ June ٢٠٠٢.

(١) Dan Mahoney and Mariella Di Giacomo, “Flashpoint @ LANL. Gov: A Simple Smart Search Interface:” Issues in Science and Technology Librarianship (summer ٢٠٠١), available at: http://www.library.ucsb.edu/istl/OI-summer/articel٢٠html. Accessed ١٢ June ٢٠٠٢.

اختلاف اللغات المقيدة : حيث تعد اللغات المقيدة غير المتوافقة مشكلة عامة عندما يحاول المستفيدون البحث عبر ما وراء بيانات من مجالات موضوعية مختلفة أو عبر ما وراء بيانات من جهات مختلفة ، كالمكتبات والمتاحف ومتاحف التاريخ الوطني والجمعيات التاريخية، فهذه الجهات غالباً ما تستخدم لغات مقيدة متخصصة مختلفة ، كما أن بعضها الآخر قد يستخدم لغات مقيدة متخصصة جداً ؛ فمثلاً تقوم إحدى المكتبات باستخدام صيغة الاسم العام - مثلاً - "الثعلب الأحمر Red Fox "، عند تكشيف الأعمال حول هذا الموضوع ، بينما يستخدم أحد متاحف التاريخ الطبيعي صيغة الاسم العلمي " Vulpes vulpes " - لهذا الموضوع.

الوحدات في مقابل المجموعات : حيث تنشأ مشكلات معينة عندما يحاول المستفيدون الجمع بين ما وراء البيانات التي تصف الكيانات التي تأتي في وحدة واحدة unitary objects (مثل تسجيلات مارك الببليوجرافية أو تسجيلات دبلن كور) وما وراء البيانات التي يعتمد في إعدادها على خطط الوصف الأرشيفي المرمز EADs ، أو غيرها من الخطط الوصفية المعقدة متعددة المستويات. وعند مقابلة خطة معقدة من الخطط التي تعتمد على الوصف على مستوى المجموعة collection-level scheme مثل خطة الوصف الأرشيفي المرمز بصيغة عامة معتمدة على دبلن كور ، يفقد الوصف دقته وطبقيته، والعكس صحيح بمعنى أن الوصف الأرشيفي المرمز قد لا يشتمل على وصف موضوعي كاف على مستوى الوحدة أو المادة وبما لا يسمح باستنباط واصفات ذات معنى لتلك المواد أو الوحدات .

تعدد الإصدارات : إذ تثير التعقيدات الناشئة عن الإصدارات المتعددة كثيراً من الفروق الطفيفة ، كما تبرز إحدى المشكلات الشائعة في هذا الصدد بسبب المعالجة المختلفة للمستنسخات ، فإذا وصفت إحدى الصور المرقمنة لأحد المباني مثلاً ، فلا شك أن بعض عناصر ما وراء بيانات - كعنصري "المنشئ" و"تاريخ الإنشاء" - سوف تختلف باختلاف صيغة هذه الصورة وما إذا كان الوصف سيركز على المبنى نفسه أم على الصورة الضوئية، أم على ملف الصورة.

تعدد اللغات: فرغم أن وحدة اللغة تعد هي السمة الغالبة على المجموعات الفردية ، إلا أن شبكة الإنترنت تعد ذات طبيعة دولية ، وتثير عملية البحث البيني عبر النظم المتعددة في بيئة الإنترنت بشكل متزايد مشكلات تعدد اللغات؛ لذلك فإن الاتجاه التقليدي يجنح نحو استخدام اللغات المقيدة لأغراض الوصول أو الإتاحة ، واستخدام المكانز متعددة اللغات لأغراض الترجمة أو التحويل ، وقد تكون هناك حاجة أيضاً لمسجلات ما وراء البيانات متعددة اللغات لتساعد على عملية المقاربة بين أسماء العناصر.

READINGS :

- Arms, William Y., et al. "A Spectrum of Interoperability: The Site for Science Prototype for the NSDL." D-Lib Magazine ٨, no. ١ (January ٢٠٠٢). Available at http://www.dlib.org/dlib/januaryO٢/arms/O١ arms.html.
  How researchers approached metadata interoperability for the National Science Digital Library.

- Coyle, Karen. "The Virtual Union Catalog: A Comparative Study." D-Lib Magazine ٦, no. ٣ (March ٢٠٠٠). Available at http://www.dlib.org/dlib/marchOO/coyle/٠٣coyle.html.
  This article compares the implementation of a virtual union catalog using Z٣٩,٥٠ with the MELVYL physical union catalog.

- Lynch, Clifford A. "Metadata Harvesting and the Open Archives Initiative." ARL Bimonthly Report ٢١٧ (August ٢٠٠ ١). Available at http://www.arl.org/newsltr/٢١٧/mhp.html.
  An explanation of the Open Archives Metadata Harvesting Protocol, how it was developed, and why it is important.

- Moen, William E., and Teresa Lepchenske. Z٣٩,٥٠: Selected List of Resources. Available at http:// www.unt.edu/wmoen/Z٣ ٩ ٥٠(BasicZReferellces.htm.
  This bibliography is not entirely up-to-date but complements the bibliography on the Z٣٩,٥٠ Maintenance Agency website.

- St. Pierre, Margaret, and William P. laPlant, Jr. Issues in Crosswalking Content Metadata Standards. ١٩٩٨. Available at http://www.niso.org/press/whitepapers/crsswalk.html.
  A discussion of the incompatibilities in the representation of different metadata schemes and how these lead to problems in developing crosswalks.

- Z٣٩,٥٠: A Primer on the Protocol. Bethesda, Md.: NISO Press, ٢٠٠٢. Available at http://www. niso.org/ standards/ resources/Z٣ ٩٥٠ _primer. pdf.

Extensive web-accessible information is available on Z٣٩,٥٠, including any number of introductions and primers. This is one of the most recently published.

- Z٣٩,٥٠ International Standard Maintenance Agency: Library of Congress Network Development and MARC Standards Office (home page of the Z٣٩,٥٠ Maintenance Agency). Available at http://www.loc.gov/z٣٩٥٠/agency/.

The official maintenance site for Z٣٩,٥٠, this page links to the Library of Congress Gateway to Z٣٩,٥٠-accessible catalogs as well as to a bibliography, a list of software products that support Z٣٩,٥٠, and other related resources..

الفصل الخامس

ما وراء البيانات والوب

يعد تضمين ما وراء البيانات في الكيان الرقمي الذي تصفه من أشهر السبل للجمع بين ما وراء البيانات والمحتوى الذي يمكن الوصول إليه عن طريق الوب، فإذا كان الكيان عبارة عن وثيقة مهيأة في صيغة لغة ترميز النصوص الفائقة HTML ، فيمكن تضمينها ما وراء البيانات الخاصة بها عن طريق عناصر ما وراء البيانات أو ما تعرف بعناصر الميتا <meta> " " - كما أشرنا إلى ذلك في الفصل الثاني - بحيث يمكن بعد ذلك استخلاص ما وراء البيانات وتكشيفها عن طريق محركات البحث على الإنترنت.

محركات بحث الإنترنت :

تستخدم محركات بحث الإنترنت برامج تسمى العناكب spiders ، أو زواحف الوب webcrawlers ، في بحث محتوى الوب وجمعه وتكشيفه . ويستخدم محرك " جوجل Google " المعروف - على سبيل المثال - برنامجاً يسمى " جوجل بوت Googlebot " ؛ بينما يستخدم محرك " ألتافيستا AltaVista " برنامجاً يسمى " سكوتر Scooter " . وقد تجعل إحدى الجهات التي تشغل برنامجاً عنكبوتياً لغرض معين كجمع الصفحات المتاحة على شبكة الإنترانت Intranet الخاصة بإحدى الجهات ، أو بجمع كل المواقع الحكومية في إحدى الولايات = هذا البرنامج العنكبوتي مقتصراً على زيارة عناوين معينة دون غيرها ، ولكن الغالب هو أنه بمجرد تغذية العناكب بقائمة أولية من عناوين الوب المطلوب زيارتها ، فإنها تقوم بتوسعة القائمة بشكل مستمر عن طريق استخلاص وتتبع المحددات الموحدة لمواقع المصادر URLs التي تجدها في الصفحات التي جمعتها، كما تقوم العناكب

كذلك بشكل دوري بإعادة زيارة المواقع حتى تتمكن من جمع المحتويات الجديدة والمحتويات المعدلة. ولمحركات البحث الرئيسية لوغاريتمات معقدة تحدد مواعيد إعادة زيارة المواقع ، مع وضعها في الحسبان المعدل الذي يتم به تحديث صفحات الوب.

وعندما يجد البرنامج العنكبوتي صفحة تشتمل على ما يستحق جمعه ، يقوم بإنشاء واختزان ملخص لتلك الصفحة ، ثم يقوم بجمع الكلمات التي توجد في الصفحة واختزانها في قاعدة بيانات كشفية an index database ، وعندما يقوم المستفيد بإجراء بحث في الإنترنت باستخدام أحد محركات البحث كـ" جوجل " أو غيره من محركات البحث ، فإنه في واقع الأمر يبحث في تلك القاعدة الكشفية، وليس في صفحات الوب المصدرية نفسها ، وما يعرض عليه هو الملخص الذي كان قد اختزن من قبل ؛ ولهذا السبب لا نجد كثيراً من المداخل التي تسترجعها محركات البحث محدثة، كما قد تسترجع في بعض الأحيان صفحات لم يعد لها وجود فعلي على أرض الواقع .

وتجدر الإشارة إلى أن صفحات الوب لا تكون جميعها قابلة للجمع من قبل محركات البحث ؛ حيث يمكن لمديري هذه الصفحات الوب webmasters اتخاذ حيل تمنع الآخرين من جمع المواقع التي يديرونها، وذلك عن طريق إدراج ملف يسمى ملف الروبوت "robots.txt" في الدليل الخاص بالخادم webserver directory ، ومن الممكن أيضاً منع تكشيف صفحة بعينها عن طريق تضمينها تاج الميتا <meta> في الصيغة التالية: <META NAME="ROBOTS" CONTENT="NOINDEX">. ومن الملاحظ أن معظم

محركات البحث الرسمية تدعم طريقة ملف الروبوت "robots.txt" لمنع التكشيف، وفي مقابل ذلك لا يدعم سوى القليل منها طريقة تاج الميتا <meta> في هذا الصدد.

إن معظم محركات البحث لديها طريقة ما لتصنيف أو ترتيب نتائج الاسترجاع ترتيباً طبقياً ranking بحيث تعرض المداخل حسب درجة صلتها بالموضوع أولاً ؛ وعادة ما يضع اللوغاريتم في الاعتبار موقع كلمات البحث في الصفحات المكشفة ، بحيث تأخذ مصطلحات البحث التي تقع قرب أعلى الصفحة أو داخل تاج العنوان <title> في صفحة الوب وزناً أكبر من حيث صلتها بموضوع البحث . كما تستخدم محركات البحث أيضاً معدل تكرار المصطلحات في حساب مدى الصلة بالموضوع ، مع أنه قد يحدث أن تسقط الصفحة بأكملها من الكشاف إذا ما تكررت المصطلحات بشكل كبير ، كما سنعرف عند مناقشة عملية " الإثقال المتعمد للكلمات في صفحات الوب spamming " في فقرة لاحقة . وهناك عاملان آخران لهما تأثيرهما في عملية الترتيب الطبقي حسب درجة الصلة بالموضوع relevancy rankings ، وهما : محلل الرابط link analysis ، ومقياس تتبع الروابط المسترجعة clickthrough measurement ؛ حيث يقوم محلل الرابط بالتعرف إلى مرات الوصول للصفحات من قبل الصفحات الأخرى عن طريق الروابط الفائقة، وبناء على ذلك تأخذ أكثر الصفحات ارتباطاً بغيرها من الصفحات الأخرى more frequently linked-to pages وزناً أكبر ، وكأن ذلك أشبه بنسخة وب من كشاف الاستشهادات المرجعية، وتلك هي الكيفية التي يرتب بها محرك البحث " جوجل " المواد التي يسترجعها . أما

مقياس تتبع الروابط المسترجعة فيقوم بتتبع كيفية استجابة المستفيدين لقوائم النتائج المسترجعة hits التي يسترجعها محرك البحث استجابة لاستفساراتهم ؛ بمعنى أن تتبوأ الصفحات التي يتم اختيارها بشكل متكرر مكانة متقدمة في الترتيب الطبقي .

إن محركات بحث الإنترنت لهي بالفعل أدوات رائعة ؛ إذ يستطيع محرك البحث الواحد تكشيف بليون صفحة وب أو أكثر من ذلك ، وهو بذلك يكشف للباحثين عن كميات هائلة من محتويات صفحات الوب ، ومع ذلك يستطيع المستفيد أن يتوقع استرجاع مجموعات أكبر من ذلك؛ فالبحث في جوجل - مثلاً- عن موضوع " محركات بحث الإنترنت" يكشف عن حوالي مليوني مدخل، ومن هنا يكتسي الترتيب الطبقي وفقاً لدرجة الصلة أهمية كبيرة جداً في هذا الصدد، كما أن استخدام ما وراء البيانات في شكل تيجان الميتا <meta> يرفع من معدلات دقة المواد المسترجعة وترتيب النتائج وفقاً لدرجة الصلة بالموضوع . وإذا كان المؤلفون يستطيعون نظرياً تقديم معلومات دقيقة عن محتوى صفحاتهم على الوب في تيجان الميتا <meta> ، بحيث تستطيع محركات البحث استخدام محتوى تلك التيجان في عمليتي التكشف والترتيب، فإن استخدام تيجان الميتا <meta> لتكشيف الصفحات يختلف - عملياً - من محرك بحث لآخر. وتقوم غالبية محركات البحث الكبرى في الوقت الحالي بتكشيف المصطلحات الموجودة في حقل الوصف ، وبعضها يقوم بتكشيف المصطلحات الموجودة في حقل الكلمات الأساسية ، ولكن القليل منها ما يستخدم المحتوى الموجود في هذين الحقلين معاً لحساب الترتيب الطبقي للصفحات وفقاً لمدى الصلة ، ذلك لأنها لا تعول

كثيراً على ما وراء البيانات التي يقدمها المؤلفون . وهنالك منافسة محمومة بين مواقع الوب التجارية فيما يتصل بالترتيب الطبقي عالي المستوى عند الاسترجاع ، كما أن هناك كذلك ممارسات غير سوية يشيع استخدامها في هذا الصدد ، من هذه الممارسات " الإثقال المتعمد للكلمات في صفحات الوب spamming " ، أي إثقال الصفحة أو تحميلها بكم كبير من الكلمات الأساسية من أجل رفع معدلات استرجاعها من قبل محركات البحث . ولإخفاء هذه الكلمات الأساسية عن المستفيد يتم إدخالها عادة بلون خلفية الصفحة نفسه، أو كمحتوى لتيجان الميتا <meta>. مثال ذلك أن يقوم أحد موزعي قطع غيار السيارات بتكرار كلمات معينة مثل : كلمة " سيارات automotive " مئات المرات محاولة منه لرفع ترتيب صفحته الرئيسية. والأسوأ من ذلك أن يلجأ أحد المواقع الإباحية إلى تضمين كلمات معينة مثل "لوازم سيارات automotive supplies " بين كلماته بحيث تسترجع هذه المواقع مع نتائج البحث عن قطع غيار السيارات. ولهذا السبب يتجاهل عديد من محركات البحث تيجان الميتا <meta> عند الترتيب الطبقي ، بل إن بعضها - مثل جوجل - يتجاهلها كذلك عند عملية التكشيف.

ومما يؤسف له أن سوء استخدام ما وراء البيانات من قبل أقلية غير سوية قد يحرم الأغلبية استخدامها استخداماً فعالاً، لذلك وضعت مفاهيم بعيدة المدى لحل هذه المشكلة ، منها تطوير آليات للتحقق من دقة ما وراء البيانات وموثوقيتها يتم بمقتضاها تضمين خطط ما وراء البيانات أماكن محددة لتسجيل ما وراء بيانات عن ما وراء البيانات، مثال ذلك البيانات الخاصة بهوية المنشئ ومدى اكتمال البيانات المسجلة ودقتها، بناء على ذلك

يمكن لأرقام التدقيق المكودة encrypted Checksums المحسوبة في الوثائق وما وراء بياناتها المتضمنة بها أن تبين إذا ما كانت هناك تغييرات قد أجريت على تلك الوثائق منذ إنشائها أم لا . ومن المحتمل أن يتطلب الأمر في النهاية استخدام التوقيعات الرقمية لإنشاء ما وراء البيانات الموثوق فيها ، والتوقيعات الرقمية هذه هي بيانات تستخدم للتحقق من هوية المنشئ ومحتوى الملف باستخدام التشفير بالمفتاح العام public-key encryption . وهو أمر له أهمية خاصة في تطبيقات معينة مثل التجارة الإلكترونية وإدارة الحقوق ، حيث تعد صحة ما وراء البيانات وتكاملها من الأمور الحيوية في هذا الصدد ؛ لذلك تقوم لجنة التوقيعات الرقمية التابعة لائتلاف الشبكة العالمية W٣C حالياً بإقرار التوقيعات الرقمية بالنسبة لما وراء البيانات وكذا للوثائق المعدة بلغة الترميز الموسعة[١].

محركات البحث محددة المجال **(Domain-specific)** :

لا تتمكن محركات البحث الرئيسية العامة مثل : جوجل، وألتا فيستا، وهوت بوت ، وإجزايت إلا من استخدام الحد الأدنى من ما وراء البيانات المدمجة في تيجان الميتا <meta>، وهذا مما يؤسف له لأن هذه التيجان تتيح الفرصة للمصطلحات أن تحدد كعناصر بيانات مخصصة، كما هو الحال مع عنصري المؤلف أو الموضوع ، ومن ثم فعند إجراء بحث باستخدام محرك البحث " جوجل " أو ما شابهه من محركات البحث العامة، لا يستطيع المستفيد سوى إجراء البحث إلا بكلمات أساسية عامة فقط، كما لا يستطيع

(١) W٣C Technology and Society Domain, XML Digital Signatures Activity Statement, available at: http://www.w٣.org/Signature/Activity.html. Accessed ١٣ June ٢٠٠٢.

تحديد مجال البحث بالمؤلف أو بالعنوان أو بالموضوع مثلاً، وهذا من شأنه أن يؤدي إلى ارتفاع معدل نتائج الاسترجاع مع انخفاض معدل دقتها أو صلتها بالموضوع .

ومع ذلك، فليس هناك - عمومًا - قيود على استخدام تقنية محرك البحث مع ما وراء البيانات البنائية structured المتضمنة في صفحات الوب المرمزة بلغة ترميز النص الفائق . ويمكن للجهات التي ترغب في تكشيف مجموعات معينة من المواقع الاختيار من بين محركات البحث التجارية ومحركات البحث غير التجارية التي يمكن تهيئتها لتكشيف محتوى عناصر الميتا <meta>. ومن أمثلة هذه المحركات : Ulraseek ، و Berkeleys Swish-E ، و Microsoft's Index Server ، و Blue Angel Technologies' MetaSatr ، و Verity ، و LiveLink Search and spider .

ومن الأمثلة الجيدة على استخدام الهيئات لما وراء البيانات البنائية مع محركات البحث ما تقوم به حكومات بعض الولايات، فهناك عدد من الولايات لديها مبادرات حكومية موسعة لتيسير الإتاحة العامة للمعلومات الحكومية عن طريق تطوير بوابات مركزية إلى مصادر المعلومات، ومن الأساليب المتبعة في هذا الصدد ما تقوم به ولايات واشنطن وإلينوي ويوتا وغيرها من الولايات الأخرى من تشجيع مؤسسات الولاية على دمج ما وراء البيانات في صفحاتهم على الوب باستخدام مجموعة أساسية من عناصر البيانات الخاصة بالولاية، أو ذات الطابع الحكومي ، والتي يعتمد في اختيارها على مواصفة خدمة محدد المعلومات الحكومية (Government Information Locator Service "GILS") (انظر الفصل ١١)؛ كما قامت

بعض الولايات بتطوير مولدات ما وراء بيانات قائمة على الوب web-based generators لتيسر على موظفي الهيئات إنشاء تيجان الميتا المناسبة. وتوفر هذه الواجهات نماذج وب web forms لإدخال البيانات بحيث تُضمن أو تُربط مع اللغات المقيدة التي يلزم استخدامها في حقول معينة، وعندما يكتمل إدخال البيانات تقوم البرامج ببناء تيجان الميتا المناسبة وعرضها، والتي يمكن بعد ذلك قصها ولصقها في سياق لغة ترميز النص الفائق الخاصة بصفحة الوب(١).

أما البحث في مواقع الهيئات المختلفة فيقوم به جامع تديره الولاية a state-run harvester مهيأ للتعرف إلى تيجان الميتا التي تدعمها الولاية ويقوم بتكشيفها . ومن الملاحظ أن استخدام ما وراء البيانات البنائية يتيح للعامة إمكانية البحث عن أنواع معينة من البيانات ، كما يتيح لهم كذلك إمكانية بحث النصوص الكاملة لصفحات الوب ، ومن أمثلة ذلك واجهة ولاية واشنطن التي تتيح البحث بسبعة عشر عنصرًا محددًا من عناصر ما وراء البيانات، مثل : عنوان الوثيقة ، والقسم الذي أنتجها ، والبرنامج الخاص بالجهة، والكلمات الموضوعية الأساسية بالوثيقة .

ما وراء البيانات والصيغ في غير لغة ترميز النص الفائق NON-HTML Formats :

تقوم معظم محركات البحث على الإنترنت فقط بتكشيف الصفحات في

---

(١) As an example of a GILS-compliant metadata generator, see the Utah HTML Metatag Builder at: http:/www.utah.org/GILS/uttagbuilder.htm

صيغة لغة ترميز النص الفائق الساكنة static HTML ، ولكن هناك كماً هائلاً من المحتوى المتاح على الإنترنت متاح بصيغ غير تلك الصيغة الساكنة ، وهو ما يعرف بالوب الخفية the hidden web أو الوب العميقة the deep web ، ومثل هذا المحتوى يشتمل على قواعد بيانات متخصصة مثل : فهارس المكتبات ، وبيانات الإحصاء السكاني ، وبنوك الأخبار ، التي لا تتوفر على الوب إلا كصفحات مهيأة بشكل ديناميكي استجابة لتساؤلات البحث ، كما أنها تشتمل كذلك على صور ومواد سمعية ومواد مرئية وغيرها من الملفات غير النصية الأخرى . ويقدر حجم الوب الخفية بنحو خمسمائة ضعف لحجم الوب الظاهرة - عكس الخفية - التي تقوم محركات البحث بتكشيفها(١)؛ ونظراً لاستحالة الوصول العالمي للوب الخفية، عادة ما تكون معظم اتجاهات طرق الوصول إلى تلك المواد متخصصة موضوعياً subject-specific ، ومن ثم يتم تكوين قوائم بالمصادر القابلة للبحث، كما هو الحال مع بوابات المكتبات ، كما يتم تهيئة بعض أدوات سطح المكتب لبحث مجموعات من المصادر بشكل تزامني، كما يتم في بعض المجالات ، خاصة المجالات ذات الاهتمامات التجارية، تجميع محتويات مختارة في مستودعات خاصة بها.

إن استخدام ما وراء البيانات يعد توجهاً آخر لإتاحة الوصول لبعض محتويات الوب الخفية ، حيث يمكن تضمين ما وراء البيانات في الكيانات الرقمية في غير صيغة لغة ترميز النصوص الفائقة. وهذا ما نجده في

---

(١) Michael K. Bergman, "The Deep Web: Surfacing Hidden Value;' Journal of Electronic Publishing ٧,no ١ (August ٢٠٠١}, available at

http://www.press.umich.edu/jep/٠٧٠١/bergman.html. Accessed ١٣ June ٢٠٠٢.

الإصدارة رقم ٥,٠ وما بعدها من برنامج " أدوبي أكروبات " وغيره من منتجات شركة أدوبي مثل : InDesign وIllustrator ، التي تدعم برنامج أدوبي لما وراء البيانات الموسعة (XMP) (١) Adobe's extensible Metadata Platform، والتي تسمح بدورها بدمج ما وراء البيانات النصية في وثائق بي دي إف PDF. كما تسمح التقنية بدمج ملفات لغة الترميز الموسعة، المسماة بدفعات برنامج ما وراء البيانات الموسعة XMP packets في المستندات، وتستطيع ملفات برنامج ما وراء البيانات الموسعة هذه القيام بوصف الوثائق بأكملها أو أجزائها الأساسية، حيث يمكن أن يتضمن أحد التقارير الذي يتضمن صورة - على سبيل المثال - ملفاً واحداً من ملفات برنامج ما وراء البيانات الموسعة للنسخة النصية من هذا التقرير، وملفاً آخر للتقرير ولكن في شكل صورة وليس نصاً.

وتتكون ملفات برنامج ما وراء البيانات الموسعة من ترويسة header وتذييلtrailer تقع بينهما ما وراء البيانات. ويجب التعبير عن ما وراء البيانات نفسها في صيغة لغة الترميز الموسعة المتوافقة مع إطار وصف المصدر RDF-compliant XML . وقد حددت شركة أدوبي بعض الخطط الافتراضية لبرنامج ما وراء البيانات الموسعة لأنواع متعددة من المواد، مع أنه يمكن تحديد أي خطة طالما أنها تكون متوافقة مع مواصفة برنامج ما وراء البيانات الموسعة، وللاستفادة من ما وراء البيانات يجب على برامج

(١) Platform بيئة التشغيل : الآليات أو نظام التشغيل المناسب لتشغيل إحدى البرمجيات ويعرف أحيانا ببيئة تشغيل الكمبيوتر - أما في الاستخدام اليومي فيعني نوع الكمبيوتر أو خطة التشغيل المستخدمة [المترجم].

التكشيف القيام بمسح الوثائق بحثاً عن ملفات برنامج ما وراء البيانات الموسعة المتضمنة بها ، ثم استخلاص ما وراء البيانات لتكشيفها.

ويمكن للصيغ غير النصية أن تشتمل على ما وراء بيانات نصية في ملفات الترويسة ، كما يمكن أن تحتوي ملفات الصور التي تعد وفقاً لصيغة ملفات الصور الرقمية المعروفة بصيغة تيف TIFF على ما وراء بيانات في ترويسة هذه الصيغة - أي صيغة تيف - حيث تشير الترويسة إلى قائمة أدلة ملف الصور Image File Directories (IFDs) التي يتضمن كل منها حتماً قوائم بالتيجان، ولكل تاج منها قيمة رقمية، ونوع البيانات، ومحدد موضع البايت byte offset [١]، الذي يحدد مكان وجود البيانات (وهي صيغة بنائية مشابهة لبنية صيغة مارك)، ومن الملاحظ أن معظم معلومات الترويسة المعرفة مسبقاً هي عبارة عن ما وراء بيانات فنية توضح بالتفصيل الخصائص المادية للملف مثل : عدد البتات Bits في العينة، ومعدل الضغط، ونوع الملف . هذا مع وجود تيجان خاصة باسم الوثيقة، ووصف الصورة ، وتاريخ الإنشاء والوقت الذي تم فيه، واسم الفنان. أما عناصر البيانات المحلية فيمكن تعريفها بتيجان مرقمة [تبدأ] من الرقم ٣٢٧٦٨ فأعلى.

وقد تحتوي صيغ الصور الأخرى كذلك على ما وراء بيانات متضمنة بها، إذ تحتوي ترويسة الطريقة المعيارية لضغط الصور الفوتوغرافية الرقمية - المعروفة بصيغة "جي بيج JPEG " ـ على حقول خاصة

(١) offset تعني في أساليب العنونة النسبية رقمًا يشير إلى مدى بعد مادة معينة عن نقطة البداية[المترجم] .

بالتطبيقات (تسمى علامات آي بي بي APP markers) مع إدخال بيانات من برامج تطبيقية مختلفة. حيث يستطيع برنامج (فوتوشوب) مثلاً إدخال ما وراء بيانات بنائية في العلامة "APP١٣ " . ومع أن البرامج تستطيع استخلاص محتوى ما وراء البيانات من ترويسات الصور لتكشيفها ، فإن هذا ما لا يفعله أي من محركات البحث الكبرى. أما المحركات التي تقوم بتكشيف الصور فتستخدم مفاتيح clues نصية خارجية للصور نفسها ، بما في ذلك محتوى خاصية "alt" لتاج الصورة (<img>) في صيغة لغة ترميز النصوص الفائقة ، أو النص الذي يظهر قرب الصورة على الصفحة ، أو اسم ملف الصورة نفسها.

أما الصيغ الصوتية مثل صيغتي WAV و MP٣ فتسمح بتسجيل بعض عناصر ما وراء البيانات الوصفية في الترويسات ، فمثلاً يحتوي التاج الصوتي المسمى MPEG Audio Tag ID٣vl على مواقع ثابتة لكل من : اسم التسجيل الصوتي ، واسم الفنان ، والألبوم ، والسنة ، والتعليق ، ونوعية التسجيل، والتي تؤخذ من قائمة استنادية ذات قيم مشفرة. كذلك تقدم كل من صيغتي "MPEG-٧" و" MPEG-٢١ " إطاراً يتيح عمل ما وراء بيانات وصفية لمنتجات الوسائط المتعددة (انظر الفصل ١٧) ، ولكن كما هو الحال مع ترويسات الصور فإن محركات البحث لا تستخدم ما وراء البيانات هذه.

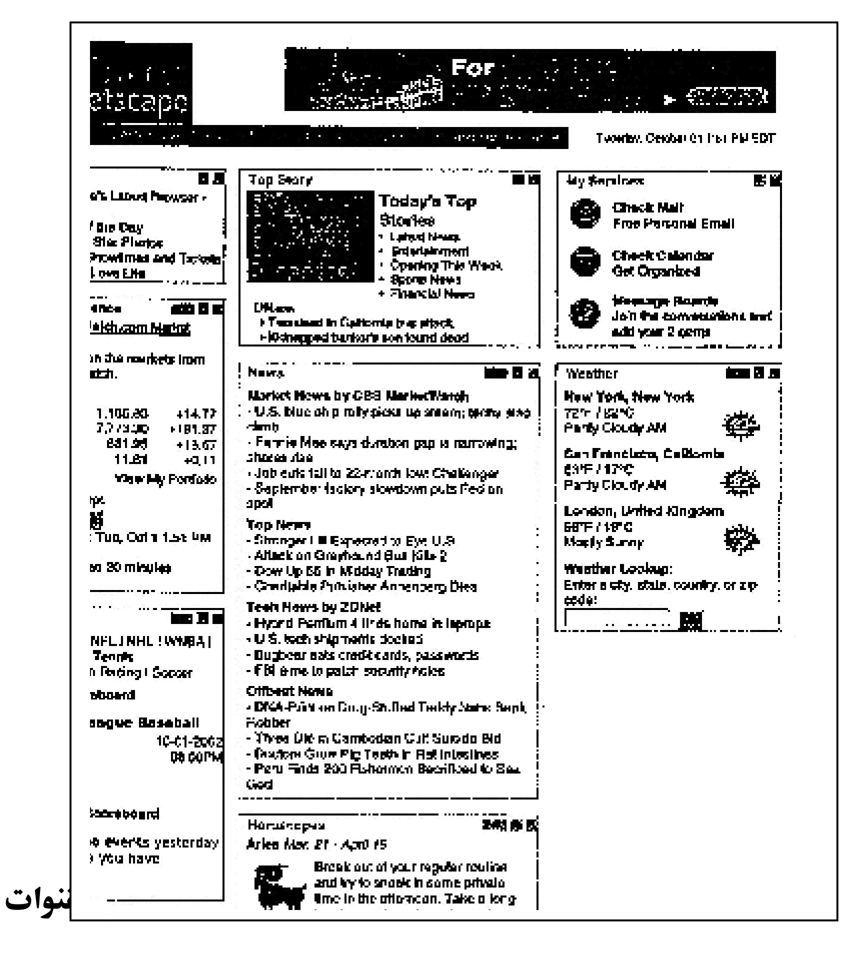
القنوات **channels** :

القناة عبارة عن موقع على الوب يمكنه أن يرسل المعلومات المحدثة

تلقائياً لعرضها بشكل فوري أو عند الطلب. وتستخدم القنوات في البث عبر الوب Webcasting (والذي يسمى أيضاً بالبث الشبكي Netcasting) ، والذي يعني التحديث مسبق الإعداد للأخبار أو أسعار الأسهم أو أية معلومات أخرى مختارة على الوب . والبث عبر الوب هو أحد أشكال تقنية الدفع push technology التي يقوم فيها خادم الوب بدفع المعلومات إلى المستفيد ، مع أن القنوات يجب أن تختار مسبقاً ، كما أن معظم عمليات الدفع تتم بناء على طلب من متصفح العميل. ويوضح الشكل رقم (١/٥) استخدام القنوات في موقع My Netscape (http://my.netscape.com) حيث يمثل كل مستطيل على الشاشة إحدى القنوات. وفي هذا التطبيق يمكن إضافة القنوات وحذفها وإعادة ترتيبها من قبل المستفيد، أي أنه يقدم خدمة معلومات مخصصة بالفعل.

وتجدر الإشارة إلى أن اختيار القنوات يتم اعتمادًا على ما وراء البيانات، وقد ظهرت العديد من الخطط التي تهدف إلى التعريف بالقنوات ومحتوياتها. وقد طورت مؤسسة آبل Apple Corporation ورعت مواصفة مبكرة في هذا الصدد، هي إطار ما وراء المحتوى the meta Content Framework(MCF) ، كما أن هناك صيغة تعريف القناة لشركة مايكروسوفت . ومن الصيغ المتداولة بشكل كبير في الوقت الحالي صيغة مختصر موقع إطار وصف المصدر RDF Site Summery (RSS) ، وهي مواصفة بسيطة إلى حد كبير ومعتمدة على لغة الترميز الموسعة . وتتواجد هذه الصيغة في عدة إصدارات ، غير أنها ليست متوافقة تماماً ؛ فهناك الإصدارة رقم ٠,٩ (

RSS ٠,٩) ، وهي أبسط صيغة ، حيث تتكون من عشرة عناصر ، وقد أصدرتها نيت سكيب Netscape في عام ١٩٩٩م لتستخدم عبر موقعها المعروف بـ My netscape .



نوات

ثم جاءت بعد هذه الإصدارة بعد وقت قصير الإصدارة رقم ٠,٩١ (Rss ٠,٩١) التي تعرف بمختصر الموقع الثري (Rich Site Summery) حيث اشتملت على أربعة عشر عنصراً إضافياً ، ثم الإصدارة ٠,٩٢ (RSS ٠,٩٢

) التي تتميز بالعديد من الخصائص الإضافية.

وفي ديسمبر من عام ٢٠٠٠م ظهرت الإصدارة رقم ١,٠ (RSS ١,٠) التي دمجت الأسماء المميزة القائمة على صيغة إطار وصف المصدر RDF Namespace في الإصدارة الأولى المبسطة (RSS ٠,٩) لتجنب المزيد من تزايد عناصر البيانات الأساسية. ومن الواضح الآن أن إصدارات كل من صيغتي ملخص الموقع الثري (RSS ٠,٩x) ، وملخص موقع إطار وصف المصدر (RSS ١.x) تسير في مسارات تطوير مستقلة بعضها عن بعض، والتي ستكون بلا شك عرضة لمزيد من التغيير في المستقبل .

وتأخذ جميع مواصفات ملخص الموقع بعين الاعتبار تعريف القنوات عن طريق وصف يبين عناوينها وخصائصها مع رابط لصفحة وب خارجية ، كما أن القنوات يمكن أن تشتمل بدورها على مواد متعددة (كرؤوس عناوين الأخبار مثلاً) والتي توصف هي الأخرى من خلال طريقة تبين عناوينها وروابطها ، كما تشمل بعض إصدارات ملخص الموقع على ما وراء بيانات إضافية خاصة بمعاملات التقدير ratings ، والجوانب اللغوية ، ومعلومات حقوق النشر والتأليف . ومع أن الإصدار رقم ١,٠ (RSS ١,٠) تقرر الحد الأدنى من عناصر البيانات الموجودة في المواصفة الأساسية ، فإنها لا تقبل إدراج النظم الفرعية الوظيفية modules - التي تشير إليها الخطة الخارجية على أنها أسماء مميزة قائمة على صيغة إطار وصف المصدر - بهدف توسعة مجموعة العناصر ، وتعد خطة دبلن كور من النظم الفرعية الوظيفية القياسية المستخدمة لتقديم ما وراء البيانات الوصفية . ويعرض الشكل رقم (٢/٥) مثالاً لإحدى القنوات المعرفة عن طريق صيغة ملخص الموقع رقم ١,٠ ، مشتملة

على عناصر مأخوذة من الاسم المميز لدبلن كور .

وإذا ما نظرنا إلى القنوات من منظور أحادي يمكن القول إنها تعد وببساطة بمثابة آليات للتواصل مع محتويات وب معينة، إلا أن ما وراء البيانات الموضوعة في صيغ كصيغة مختصر الموقع (RSS) تكتسي من منظور آخر أهمية كبيرة لدورها في استكشاف المصادر واختيارها. وتجدر الإشارة إلى أن منتجي محتوى هذه القنوات كثيراً ما يقومون ببيع الكثير منها لينشر في أكثر من قناة واحدة في الوقت نفسه، أو يبيعونه للموزعين ليعيدوا توزيعه في أكثر من قناة ، وعادة ما يكون التوجه صوب جهات التجميع أو المجمعين aggregators الذين يقومون بإعادة بيع المحتويات التي يحصلون عليها من عدد كبير من الموردين. وتتمثل القيمة المضافة المهمة لهذه الخدمات في إتاحتها القدرة على انتقاء المحتوى المناسب بشيء من الدقة بحيث يمكن تهيئة هذه المحتويات لإعادة بيعها ، بل إن بعض جهات التجميع تجيز لعملائها عمل المرشحات filters المعتادة الخاصة بهم . ولا شك أن القدرة على تقديم هذه الخدمات تعتمد على وجود ما وراء بيانات تفصيلية بما فيه الكفاية من ناحية ، ومرمزة بشكل ثابت من ناحية أخرى، ومن ثم فإن القدرة على تضمين وصف معقد إلى حد كبير للمصدر في صيغ القنوات - كصيغة ملخص الموقع هذه - له قيمته العملية والتجارية المباشرة .

الوب الدلالية **The semantic Web** :

مما يلفت الانتباه للوهلة الأولى أن معظم ما وراء البيانات التي يقوم الإنسان بإنشائها لوصف مصادر المعلومات موجهة لكي يستخدمها الإنسان نفسه، وهناك مسلمة تقول إن الشخص الذي يبحث جدياً عن المعلومات (

سواء أتم ذلك بشكل مباشر أو غير مباشر عن طريق إنشاء ملفات التفصيلية التـي تعمـل تلقائيـاً) سوف يقوم بتنقية نتائج البحث المسترجعة واختيار ما يناسبه منها ، كما سيقوم كذلك بتحليل الروابط التي تربط بين الأجزاء المختلفة للمعلومات، غير أن كثيرين من نشطاء مجتمع الوب بما فيهم تيم بيرنرز- لي Tim Berners-lee، الذي ينسب

إليه اختراع الوب عامة ، يعتقدون أن الآلات ينبغي أن تكون قادرة في المستقبل على القيام بمثل هذه المهام والمزيد منها آلياً.

وتركز لجنة الوب الدلالية التابعة لاتحاد الوب W٣C semantic web activity حالياً على تمثيل البيانات الدلالية على الوب بطرق معالجة آلياً ، وقد عبرت هذه اللجنة عن مفهومها للوب الدلالية بأنها " امتداد للوب الحالية ، تأخذ فيها المعلومات معاني أكثر وضوحاً ،وبشكل يمكّن الحاسبات والبشر من العمل سوياً وبشكل أفضل ، كما أنها تقوم على أساس إيجاد بيانات معرفة ومترابطة على الوب بطريقة تكفل لتطبيقات متنوعة استخدامها بفعالية في عمليات الاستكشاف والأتمتة والتكامل وإعادة الاستخدام(١) "، ولاتحاد الوب توضيح سابق بشأن معمارية الوب يصف فيها الوب الدلالية بأنها " تلك الوب التي تشتمل على وثائق أو أجزاء من وثائق تصف العلاقات الصريحة بين الأشياء، وتتضمن معلومات دلالية تدعم جهود المعالجة الآلية التي تقوم بها آلاتنا(٢) "(٣) . وهناك أداتان رئيسيتان يعتمد عليهما في إنشاء الوب الدلالية هما : إطار وصف المصدر RDF ، والأنطولوجيات أو مكانز المفاهيم ontologies. ويعد إطار وصف المصدر هو الأساس في هذا الصدد ؛ ذلك لأنه يقوم على أساس تمثيل ما وراء

---

(١) W٣C Technology and Society Domain, Semantic Web Activity Statement, available at http://www.w٣.org/٢٠٠١/sw/Activity. Accessed ١٤ June ٢٠٠٢.

(٢) يقصد بها الحاسبات [المترجم].

(٣) Tim Berners-Lee, Dan Connolly, and Ralph R. Swick, "Web Architecture: Describing and Exchanging Data;'W٣C Note ٧ (June ١٩٩٩), available at http://www.w٣.org/١٩٩٩/٠٤/ WebData. Accessed ١٤ June ٢٠٠٢.

البيانات كتأكيدات assertions يمكن أن تحدد بكاملها في شكل محددات موحدة لمواقع المصادر URLs ، أو ربطها بتعريفات لكل من الموضوع والكيان والسند أو من يعزا إليه التأكيد predicate of the assertion ، حيث يمكن أن يقوم إطار وصف المصدر بعمل التأكيد التالي الذي يقول أن وثيقة ما ، ولتكن " أ " (أي الموضوع) أنشأها (السند أو من يعزا إليه إعداد هذا العمل) جون سميث (أي الكيان) ، حيث يمثل الكيان كرابط بتسجيلة استناد أو بجزء من وثيقة تعرّف بالسيد سميث ، ويمثل من يعزا إليه العمل كرابط بعنصر المنشئ في مواصفة دبلن كور لما وراء البيانات ، ويمثل الموضوع كرابط بالوثيقة . ومن ثم تستطيع البرامج الحاسوبية أن تتتبع تلك الروابط للحصول على معلومات إضافية عن جون سميث أو عن منشئي الوثائق.

أما الأنطولوجيات أو مكانز المفاهيم ontologies - المكون الأساسي الثاني - فتستمد أهميتها من إتاحتها الفرصة للحاسبات لاستنباط معاني المصطلحات ، وهي عبارة عن وثائق تحدد بشكل رسمي العلاقات الدلالية بين المفاهيم ، وهي أشبه ما تكون بالمكانز ، إلا أنها تعنى بالمفاهيم وليس بالمصطلحات ، من ثم فهي تشتمل عامة على تصنيفات تعّرف بفئات الكيانات وما بينها من علاقات، وعلى قواعد الاستنباط التي تعتمد عليها في تحديد تلك العلاقات.

والفكرة الكامنة وراء الوب الدلالية هي أنه يمكن كتابة برمجيات وسيطة software agents لاستخدام التأكيدات المحددة في إطار وصف المصدر مع مكانز المفاهيم وقواعد الاستنباط ، لعمل روابط بين المعلومات غير المترابطة بوضوح أو صراحة، ولكن لسوء الحظ فعندما توصف الوب

الدلالية فإنها توصف بأمثلة مبتذلة. كما ينظر للبرامج الوسطية الآلية وكأنها السكرتيرات العاملات بالمكاتب اللواتي تقمن فقط بتنظيم المواعيد وعمل ترتيبات السفر بشكل آلي، ولكن على أية حال فإن كثيرًا من الباحثين يعتقدون بأن الوب الدلالية التي سوف تحسن كثيراً من قدرتنا على استخدام المعلومات ، سوف تثبت أنها أكثر قابلية للتحول من شبكة الوب الحالية (World Wide Web) التي حسنت بشكل كبير من قدرتنا على الحصول على المعلومات. وعلى أي حال ، يتفق الجميع على أن من أكبر التحديات لتحقيق رسالة الوب الدلالية لا يقع على عاتق تطوير مكانز المفاهيم ، أو محركات الاستنباط inference engines ، أو البرمجيات الوسيطة الذكية[١] intelligent agents فحسب ، ولكن على عاتق المؤلفين اللذين ينبغي أن يشجعوا على تقديم ما وراء بيانات ذات معنى مع مصادرهم التي يتيحونها على الوب.

READINGS :

(١) agent لها معان كثيرة ، فضلاً عن وكيل - أي الشخص الذي يعمل كوسيط بين طرفين - تعني البرنامج الذي يؤدي مهمة معينة في الخلفية وعند الانتهاء منها يبلغ المستفيد بانتهاء المهمة ، كما يعني في موديل العميل أو الخادم ذلك الجزء من النظام الذي يقوم بتجهيز المعلومات وتبادلها لصالح العميل أو الخادم ، فإذا قيل مثلاً intelligent agent فهي تشير إلى نوع من العمليات الأوتوماتيكية التي يقوم بها الوكيل بالتخاطب مع الوكلاء الآخرين لأداء بعض المهام الجماعية لصالح شخص أو مجموعة من الأشخاص . كما تعني في مجال برامج إدارة الشبكات البرنامج الذي يراقب حركة مرور الإشارات على الشبكة [المترجم].

- Ackermann, Ernest, and Karen Hartman. Searching and Researching on the Internet and the World Mde Web. ٢nd ed. Wilsonville, Ore.: Franklin, Beedle, ١٩٩٧. See also the online companion pages at http://www.webliminal.com/search/.

A comprehensive instructional guide to the use of directories, virtual libraries, search engines, and other tools for performing research on the Web. Among many other topics, this resource describes how specific search engines calculate relevance ranking, counter spamming, and perform other functions.

- Berners-Lee, Tim, James Hendler, and Ora Lassila. "The Semantic Web." Scientific American (May ٢٠٠١).

A clear explication of the vision of the Semantic Web by some of the visionaries most instrumental in its conception and design.

- Lynch, Clifford A. "Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust' Available at http://www.clir.org/pubs/reports/ pub٩٢/lynch.html.

A thought-provoking discussion of authenticity, integrity, and provenance in the networked information environment

الجزء الثاني

خـطط مـــا وراء الـبـيــانـــات

الفصل السادس

فهرسة مجموعات المكتبات

انخرط المكتبيون في وصف مصادر المعلومات منذ أن أنشئت المكتبات ، ودليل ذلك ما وجد من نظم تصنيف مستخدمة في المكتبات الآشورية والبابلية في القرن السابع قبل الميلاد(١). أما أصول القواعد الأنجلو أمريكية الحديثة فتعود إلى تجميع أنتوني بانيتزي Anthony Panizzi لقواعد فهرسة المتحف البريطاني في منتصف القرن التاسع عشر الميلادي، ثم ما تبعها من تقنينات أخرى للفهرسة ظهرت على مدار المئة والخمسين سنة التالية والتي أعدها في البداية أشخاص مرموقين مثل: تشارلز جيوليت، وتشارلز كتر، وأخيراً ما تبنته الجمعيات العلمية مثل جمعية المكتبات الأمريكية ومكتبة الكونجرس.

مبادئ الفهرسة :

أرست قواعد كتر للفهرس القاموسي Cutter's Rules for a Dictionary Catalog (التي نشرت أول مرة في عام ١٨٧٦م كقواعد للفهرس القاموسي المطبوع) مبادئ الوصف الببليوجرافي التي كان لها أثرها الواضح في جميع ما لحقها من تقنينات الفهرسة. وكانت عبارة "كتر" الشهيرة والمعروفة بمبدأ " راحة المستفيد user convenience " التي يقول فيها: "لابد أن تكون لراحة المستفيد الأولوية على راحة المفهرس"، أساساً للعديد من القواعد المستخدمة اليوم، ويرى هذا المبدأ أن الفهرس يحتاج إلى أن يرتب ترتيباً هجائياً وليس مصنفاً، كما أدت عبارته الشهيرة هذه إلى ظهور مبدأ الاستخدام الشائع common usage الذي يلزم المفهرسين أن يختاروا للتعبير

(١) John Mecalfe, Information Retrieval British and American, ١٨٧٦-١٩٧٦ (Metuchen, N.J.: Scarecrow, ١٩٧٦),٥

عن الموضوعات والأسماء المصطلحات التي تزداد احتمالات استعانة الباحثين بها في البحث أكثر من غيرها، وقد بين كتر الأهداف التي يرى من وجهة نظره أن يحققها الفهرس، ولخصها في النقاط التالية[١]:

الأهداف :

تمكين الشخص من العثور على كتاب يعرف منه أياً مما يلي:

- المؤلف.
- العنوان.
- الموضوع.
- إظهار ما تقتنيه المكتبة :

- لمؤلف معين.

- عن موضوع معين، في مجال معين من مجالات الإنتاج الفكري.

- المساعدة في اختيار الكتاب :

- حسب طبعته (ببليوجرافيا).

- حسب خصائصه (الأدبية أو الموضوعية) .

وبعد أن عبر عن هذه الأهداف قام بتحديد أدوات الفهرسة التي تساعد على تحقيقها؛ فلتمكين المستفيد مثلاً من العثور على كتاب يعرف اسم مؤلفه،

(١) Charles A. Cutter, Rules for a Dictionary Catalog, ٤th ed. (Washington, D.C.: Government Printing office, ١٩٠٤),٦.

وتوضيح ما تقتنيه المكتبة من إنتاج فكري لمؤلف معين ، على المفهرس أن يعد مدخلاً لكل مؤلف مع الإحالات اللازمة لذلك . وللمساعدة في اختيار الكتاب حسب طبعته ، على المفهرس أن يوفر بيانات الطبعة وبيانات النشر للكتب التي يصفها ، مع أي تبصرات أخرى يتطلبها الموقف.

ثم تبعت قواعد كتر تقنينات الفهرسة اللاحقة ، منها على سبيل المثال لا الحصر، قواعد الجمعية الأمريكية للمكتبات ALA ، التي نشرت في طبعات متتالية في عام ١٩٠٨م، ثم في عامي ١٩٤١م، و ١٩٤٩م، والقواعد الأنجلو أمريكية للفهرسة ، التي نشرت طبعتها الأولى في عام ١٩٦٧م، ثم الطبعة الثانية في عام ١٩٧٨م. ومنذ بداية السبعينات من القرن العشرين قام الاتحاد الدولي لجمعيات المكتبات ومعاهدها IFLA بإصدار سلسلة من قواعد التقنين الدولي للوصف الببليوجرافي ISBDs تهدف إلى توحيد معايير ممارسة الفهرسة على المستوى الدولي(١). وقد صدرت عدة تقنينات لقواعد الوصف هذه ، منها - على سبيل المثال لا الحصر - التقنين الدولي للوصف الببليوجرافي (العام) ISBD(G) ، والتقنين الدولي للوصف الببليوجرافي (للمنفردات) ISBD(M) ، والتقنين الدولي للوصف الببليوجرافي (للمسلسلات) ISBD(S). ومن الملاحظ أن الإصدارة الحالية من القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية كان آخر ما صدر منها مراجعة عام ١٩٨٨م(٢) (AACR٢R) التي اعتمدت بشكل كبير على قواعد التقنين

---

(١) IFLANE ,family of ISBDs (web page), available at http://www.ifla.org/VI/٣/nd١/isbdlist. htmt. Accessed ١٧ June ٢٠٠٢.

(٢) نشرت ثلاث إصدارات منقحة لهذه الطبعة الأساسية ، في الأعوام : ١٩٩٣م، ١٩٩٨م،٢٠٠٢م [المترجم].

الدولي للوصف الببليوجرافي ISBDs.

ومع أن بعض القواعد قد يكون حظها من النجاح أقل مما لغيرها ، فإنه يمكن القول بكل إنصاف إن قواعد فهرسة المكتبات قد حاولت دائماً الالتزام بالمبادئ الأساسية، بما في ذلك مبدأ "راحة المستفيد"، كما حاولت دائماً أن تسعى لتحقيق الأهداف المثلى للفهرس. وقد انعكست أهداف كتر بعد ذلك بأكثر من مائة عام على المهام الثلاث الأولى من "مهام المستفيد user tasks " الأربعة التي حددها الاتحاد الدولي لجمعيات المكتبات ومعاهدها في سياق المتطلبات الوظيفية للتسجيلات الببلوجرافية، والتي جاءت على النحو التالي:

- البحث عن كيانات تتفق مع معيار البحث الذي يقرره المستفيد (بمعنى تحديد مكان كيان واحد أو مجموعة من الكيانات في أحد الملفات أو إحدى قواعد البيانات استجابة لبحث باستخدام خاصية الكيان أو علاقته بغيره من الكيانات).

- تعريف الكيان أو تحقيق هويته (بمعنى التأكيد على أن الكيان الموصوف يرتبط مع الكيان الذي يبحث عنه، أو التمييز بين كيانين أو أكثر تجمع بينهما خصائص متشابهة).

- اختيار كيان يتفق واحتياجات المستفيد (بمعنى اختيار كيان ما يلبي احتياجات المستفيد من حيث المحتوى والشكل المادي وغير ذلك من خصائص ، أو رفض كيان ما لكونه لا يتفق واحتياجات المستفيد).

- الحصول على الكيان الموصوف أو إتاحة إمكانية الوصول إليه

(بمعنـى الحـصول علـى كيان مـا عـن طـريق الـشراء أو الإعـارة أو غيـرهما مـن سـبل الاقتـناء ، أو الوصـول إليه بطـريقة إلكتـرونية مـن خلال الاتصال بالإنترنت عن طريق حاسب بعيد)[1].

قواعد الفهرسة ومواصفاتها :

يحتاج المفهرس اليوم إلى عدد كبير من القواعد والموجزات الإرشادية والأدلـة التوثيقية. وتعتبر الإصدارة المراجعة من القواعد الأنجلو الأمريكية للفهرسـة لعـام ١٩٩٨م (AACR٢R) هـي المرشـد الرئيـسي لإنـشاء محـتوى الفهرسـة الوصـفية ، بمـا فـي ذلـك صـياغة نقاط الإتاحة ، غير أن ممارسات الفهرسـة التـي تتبـناها مكتبة الكونجرس قد تختلف إلى حد ما عن حرفية تلك الإصـدارة المـراجعة الأخيـرة ، وقد تكفل برصـد هذه الاختلافات ما يعرف بتفسيرات مكتـبة الكونجـرس لقـواعد الفهرسـة Library of congress Rule Interpretations "LCRIs" الذي تنشره المكتبة دورياً . ولأن مكتبة الكونجرس تعد هـي أكبر مصدر لتسجيلات مارك٢١ ، ولأهمية التناغم الذي ينبغي أن يحكم ممارسات الفهرسة التعاونية ، فإن معظم أقسام الفهرسة تتبع تفسيرات مكتـبة الكونجـرس للقـواعد ، وهـناك تـوجه لتشجيع الاعتماد عليها عند إعداد التسجيلات لإضـافتها لقاعدة بيانات مركز الحاسب الآلي للمكتبات على الخط

(١) International Federation of Library Associations and Institutions, Functional Requirements for Bibliographic Records: Final Report, UBCIM Publications, new series, vol. ١٩ (Munich: K. G. Sa r, ١٩٩٨), available at

ttp://www.ifla.org/VII/s١٣/frbr/frbr.htm.Accessed ١٧ June ٢٠٠٢.

المباشر المعروف بـ " أو سي إل سي OCLC.

وتتطلب فهرسة الأسماء والموضوعات مجموعة قواعد إضافية ، وأدلة توثيقية مساندة ، وملفات استناد ؛ حيث يتم إنشاء مداخل الأسماء أو رؤوس الموضوعات للأسماء name headings اعتماداً على القواعد التي يتضمنها الجزء الثاني من القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية، وتفسيرات مكتبة الكونجرس لقواعد الفهرسة. كما يمكن البحث عن مداخل الأسماء هذه في الملف الاستنادي الوطني للأسماء NAF الذي تتيحه معظم المرافق الببليوجرافية على الخط المباشر. كما يمكن للمكتبات أن تسهم بإضافة مداخل جديدة في الملف الاستنادي الوطني للأسماء من خلال ما يعرف بـ "المكون الاستنادي للأسماء " the name authority component "NACO" التابع لبرنامج الفهرسة التعاونية، الذي ينفرد بمجموعة خاصة به من السياسات والأدلة التوثيقية(١).

أما الفهرسة الموضوعية فلم تتطرق إليها القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية ؛ ولذلك تتبع المكتبات التي تستخدم خطة رؤوس موضوعات مكتبة الكونجرس LCSH الدليل الإرشادي لمكتبة الكونجرس للفهرسة الموضوعية LC's subject cataloging manual، كما تقوم بتحديد رؤوس الموضوعات من خلال ما يعرف بـ " الكتب الحمراء(٢) the red

(١) Program for Cooperative Cataloging NACO home page, available at http://www.loc.gov/ catdir/pc /naco.html. Accessed ١٧ June ٢٠٠٢.

(٢) تسمية تشتهر بها قوائم قائمة رؤوس موضوعات مكتبة الكونجرس في مجلداتها المتعددة [المترجم].

books " وتحديثاتها المباشرة الأسبوعية(١)، كما لخطط رؤوس الموضوعات الأخرى مجموعتها الخاصة من الأدلة التوثيقية ، كما هو الحال مع خطط التصنيف المختلفة.

وتشير كل من القواعد الأنجلو أمريكية للفهرسة في طبعتها المراجعة (AACR٢R) وخطط الفهرسة الموضوعية إلى محتوى الوصف الببليوجرافي فقط ، أما قواعد تمثيل هذا المحتوى في صيغة الفهرسة المقروءة آلياً أو ما يعرف بمارك فتتكفل بها صيغة مارك٢١ للبيانات الببليوجرافية MARC٢١ Format for Bibliographic Data التي تتبنى نشرها وتحديثها مكتبة الكونجرس ، وهناك نسخة مختصرة منها متوافرة على الإنترنت وتحتوي على بيانات وصف وأمثلة محدودة ، تعرف بصيغة مارك٢١ المختصرة للبيانات الببلوجرافية MARC٢١ Concise Format for Bibliographic Data (٢). وتجدر الإشارة إلى أن لصيغة مارك٢١ متطلبات للمحتوى خاصة بها لا يتم تغطيتها من جانب القواعد الأنجلو أمريكية للفهرسة ، مثال ذلك : ترميز حقول الضبط (كالحقل رقم ٠٠٧ ، والحقل رقم ٠٠٨)، وحقلي المقتنيات Holdings والمواقع locations ، وحقول التبصرات التي لم تغطها قواعد الفهرسة ، كما تشمل وثائق مارك٢١ أساليب التوثيق الرسمي

---

(٣) Bibliographic and order information for all these publications is available from the Cataloging Distribution Service. Library of Congress, Cataloging Distribution Service, Bibliographic Products and Services (home page), available at http://lcweb.loc.gov/cds/. Accessed ١٧ June ٢٠٠٢.

(١) Library f Congress, Network Development and MARC Standards Office, MARC٢١ Concise format for Bibliographic Data, ٢٠٠١ Concise Edition, available at http://lcweb. loc.gov/marc/bibliographic/. Accessed ١٧ June ٢٠٠٢.

الخاصة بكل من المحتوى وترميز عناصر هذه البيانات. و تلحق بتوصيفات مارك٢١ نفسها قوائم مقننة مستقلة stand-alone code lists ، فهناك قوائم استناد للقيم المرمزة المطلوبة لعناصر بيانات معينة مثل: الدول واللغات والمناطق الجغرافية ، كما توفر مكتبة الكونجرس أيضاً إرشادات مستقلة لترميز حقول معينة مثل حقل الموقع والإتاحة الإلكترونية Electronic Location and Access (الحقل٨٥٦)(١).

وتجدر الإشارة إلى أن معظم عمليات الفهرسة غالبًا ما يضطلع بها أحد المرافق الببليوجرافية الوطنية (مثل مركز الحاسب الآلي للمكتبات على الخط المباشر ، أو شبكة معلومات مكتبات البحث RLIN) أو أن تتولاها نظم محلية ، ثم يتم تحميل ما ينتج عن ذلك من تسجيلات في أحد هذين المرفقين الببليوجرافيين أو كليهما ، مع ملاحظة أن لكل مرفق منهما معايير إدخال وإرشادات خاصة ، فضلاً عن مجموعات بقوائم الرموز. ويعد ما يعرف بِـ " الأشكال والمعايير الببليوجرافية bibliographic formats and standards " هو دليل الفهرسة الرئيسي لمستخدمي مرفق بيانات مركز الحاسب الآلي للمكتبات على الخط المباشر ، والذي له أفضلية الاستخدام على غيره من الأدلة التوثيقية الأخرى ؛ لأنه يجمع بين إرشادات المحتوى وإرشادات الترميز(٢).

---

(٢) Library f Congress, Network Development and MARC Standards Office, Guidelines for the Use Field ٨٥٦, available at

http://www.loc.gov/marc/٨٥٦guide.html.Accessed ١٨ June ٢٠٠٢.

(١) Bibliographic Formats and Standards, ٣rd ed. (Dublin, Ohio: OCLC Online Computer Library enter, ٢٠٠٢), available at http://www.oclc.org/oclc/bib/about.htm.

وقد يحتاج المفهرسون المعنيون بفهرسة أنواع محددة من المواد إلى أدلـة تطبـيق إضـافية؛ حـيث يتـبع مفهرسـو المسلـسلات المـشاركون فــي المشروع الوطنـي "لتحويل المسلسلات " ـ المعروف بكونسر CONSER " - مثلاً - "دليل تحرير كونسر CONSER editing guide "، و"دليل فهرسة كونـسر CONCER Cataloging Manual " ، اللـذين تكـون لهمـا أولـوية الاستخدام على غيرهما من الأدلة الأخرى(١). كذلك تنشر مكتبة الكونجرس بالتعاون مع عديد من أقسام الفهرسـة التـي تتبعها، القواعد الداخلية لمكتبة الكونجرس لفهرسـة الكتب النادرة والمطبوعات ذات الأوراق الحرة وبعض أنواع المواد الخاصة" ، حيث تختلف الممارسة اختلافاً واضحاً عن استخدام القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية في هذا الصدد .

وإذا ما وضعنا ملابسات هذه البيئة المعقدة في الاعتبار، يتضح لنا من غير شك أن هناك حاجة إلى تدريب أساسي لتطوير كفاءات القائمين على الفهرسة في المكتبات ؛ لذلك فإن إحدى لجان جمعية المكتبات الأمريكية التي قامـت بتطويـر مخطـط " العناصـر الأساسـية لبـرنامج تدريبـي للمفهرسـين المبتدئين" ، أشارت إلى أن البرنامج قد يستغرق من ستة أشهر إلى سنة(٢)،

Accessed ١٨ June ٢٠٠٢ .

(٢) CONSER Editing Guide, ١٩٩٤, available at http://www.carl.org/tlc/crs/edit٧٣٧٥.htm. Accessed ٨ June ٢٠٠٢; CONSER Cataloging Manual, available at http://www.carl.org/tlc/ crs/manl ٥٧٣.htm. Accessed ١٨ June ٢٠٠٢.

(١) Association for Library Collections and Technical Services, Training Catalogers in the Electronic Era: Essential Elements of a Training Program for Entry-Level Professional Catalog, available at

ذلك أن فهرسة المكتبات لا تتطلب مستوى عاليًا من المهارة فحسب، بل إنها عملية تستغرق وقتاً طويلاً من ساعات عمل الموظفين . وقد أشارت إحدى الدراسات المنشورة في عام ١٩٩٧م إلى أن تكلفة فهرسة أحد المنفردات وإعداده حتى وضعه على الرف تصل إلى ٤٨,١٩ دولاراً أمريكياً، كما أظهرت دراسة أخرى في جامعة ولاية أيوا نشرت عام ١٩٩٢م أن متوسط فهرسة أحد المسلسلات قد يستغرق من قسم الفهرسة نحو ١,٩٧ ساعة ، وأن متوسط فهرسة الكتاب الواحد تصل إلى ١,٣٢ ساعة(١). كما قُدرت تكلفة مكتبة الكونجرس لإنشاء تسجيلة فهرسة واحدة كاملة ما بين ٥٠ إلى ١٠٠ دولار(٢).

ونظراً للتكلفة العالية لفهرسة عمل من البداية (والتي تعرف بالفهرسة الأصلية) ، فإن المكتبات تعتمد بشكل كبير على " الفهرسة المنقولة " ؛ حيث يتم الحصول على التسجيلات المعدة مسبقاً لأحد الأعمال (النسخ)، واستخدامها كأساس تجرى عليه أية تعديلات داخلية. وعادة ما يقوم

---

http://www.ala.org/alcts/publications/educ/training.htmi. Accessed ١٨ June ٢ ٢.

(٢) Scott Ben ett, "Just in Time Scholarly Monographs;' Journal of Electronic Publishing ٤, no. ١ (September ١٩٩٨), available at http://www.press.umich.edu/jep/٠٤-٠١/bennett.htrnl. Accessed ٨ June ٢٠٠٢; Lori L. Osmus and Dilys E. Morris. "Serials Cataloging Time and Costs: Re ults of an Ongoing Study at Iowa State University," Serials Librarian ٢٢, no. ١/٢ (١٩٩٢).

(١) LC٢١: A digital Strategy for the Library of Congress (Washington, D.C.: National Academy Press, ٢OdO), ١٢٣, available at http://www٧.nationalacademies.org/cstb/pub_lc٢١.html. Accessed :١٨ June ٢٠٠٢.

بإجراءات الفهرسة المنقولة موظفون غير محترفين ، ووفقاً لإجراءات تفصيلية مكتوبة. وبالرغم من وجود عدة مصادر لعملية الفهرسة المنقولة ، إلا أن قاعدة بيانات كل من مركز الحاسب الآلي للمكتبات على الخط المباشر (أو سي إل سي) ، وشبكة معلومات المكتبات البحثية ، يعدان المصدرين الأساسيين في هذا الصدد ، مع ملاحظة أن المكتبات التي تحصل على إحدى النسخ من تسجيلات من (أو سي إل سي) تميز بين نسخة مكتبة الكونجرس - أي تسجيلات الفهرسة التي أنشأتها مكتبة الكونجرس من ناحية - ونسخة المكتبة المشاركة - أي التسجيلات التي أنشأتها إحدى المكتبات الأخرى الأعضاء في مرفق أو سي إل سي - حيث تتميز نسخ مكتبة الكونجرس عامة بالدقة الفائقة، ومن ثم فإن المكتبات تأخذها بأكبر قدر من الثقة ، ولا تتطلب سوى مراجعة طفيفة.

ويرى عدد من المهتمين بمبادرات ما وراء البيانات أنه من المعروف بداهة أن فهرسة المكتبات عملية معقدة للغاية لدرجة يصعب معها على غير المهنيين القيام بها ، كما أنها مكلفة إلى حد كبير يكون من غير العملي معه إعادة فهرسة الأشكال المختلفة للمصدر الواحد ، وتلك ملحوظات جديرة بأن تؤخذ بعين الاعتبار ؛ ولذلك تميل أقسام الفهرسة حتى متوسطة الحجم منها إلى أن يكون لديها مجموعة من الأدلة التوثيقية ، وآليات قوية للتدريب ، وإجراءات لتدفق العمل تتيح لموظفي المكتبة فرص التخصص ، هذا فضلاً عن نظم إدارة مكتبية تتوافر لها قطاعات وظيفية داعمة لعمليات الفهرسة، وإمكانات الوصول إلى نظم الفهرسة التعاونية، وأدوات برمجية أو خدمات العملاء التي تساند إجراءات الضبط الاستنادي، ذلك الجهاز الببليوجرافي

المهم الذي يندر أن يكون موجوداً خارج بيئة المكتبات.

أساسيات الفهرسة :

لا شك أنه يصعب في هذا السياق وصف إجراءات فهرسة المكتبات وصفاً كاملاً ، ولكن حسبنا فقط أن نقدم تعريفاً عاماً بهذا النمط من أنماط وصف مصادر المعلومات ، ولصعوبة الفصل كذلك بين قواعد الفهرسة والترميز بصيغة مارك ٢١ في بيئة الفهرسة اليوم ، فسوف نتناولهما معاً في الفقرات التالية .

يغطي الجزء الأول من القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية المراجعة (AACR٢R) قواعد الوصف الببليوجرافي . ويقسم الوصف في سياق هذا الجزء إلى ثمانية حقول ، يشتمل كل منها على عنصر بيانات واحد أو أكثر من عنصر قد تتفاوت معالجته حسب نوع المادة الموصوفة ، وهذه الحقول توازي الحقول أو الأجزاء الأساسية التي تتكون منها صيغة مارك ، وتأتي على النحو التالي :

- حقل العنوان وبيان المسؤولية .
- حقل الطبعة .
- حقل التفصيلات المحددة للمادة (أو نوع العمل) .
- حقل بيانات النشر والتوزيع .
- حقل الوصف المادي .
- حقل السلسلة .

- حقل التبصرات .

- حقل الترقيم الدولي وسبل الإتاحة .

ويشتمل الحقل الأول - حقل العنوان وبيان المسؤولية - على عناصر البيانات التالية : " العنوان نفسه"، مع ما يسمى بالتحديد العام للمادة(١) GMD، والعناوين الموازية ، والبيانات الأخرى للعنوان، ثم بيانات المسؤولية. وتحدد القواعد الأنجلو أمريكية للفهرسة مصدراً رئيسياً لمعلومات كل حقل من هذه الحقول (أي المصدر الذي يجب أن يؤخذ منه محتوى عنصر البيانات)، كما تحدد مصدراً بديلاً للمعلومات يكون مقبولاً عندما لا يتوافر المصدر الأساسي. ومن الملاحظ أن المصدر الرئيسي للمعلومات لعنصر العنوان نفسه يختلف حسب نوع المادة ؛ حيث يكون بالنسبة للكتاب أو المسلسل المطبوع - على سبيل المثال - هو صفحة العنوان ؛ بينما يكون بالنسبة للقرص الموسيقى المدمج القرص نفسه وأي ملصق ثابت عليه. ومع أن العنوان يجب تسجيله كما يظهر في مصدر المعلومات الرئيسي في معظم الحالات، إلا أن كلمات العنوان يجب كتابتها بالأحرف الإنجليزية الكبيرة طبقاً لقواعد التكبير الخاصة بالقواعد الأنجلو أمريكية للفهرسة في تلك الطبعة المنقحة(٢).

---

(١) يقصد به وصف نوع الوسيط التي تحمل عليه المادة ، فهل هو مطبوع أو فيلم أو تسجيل صوتي مثلاً [لمترجم].

(١) قد يكون هذا الكلام صحيحاً بالنسبة للمطبوعات بلغات معينة ولا يمكن تطبيقه مع بعض اللغات الأخرى مثل العربية [المترجم].

أمـا الـتحديد العـام للمـادة GMD فهو عنصر اختياري ، يشير إلى النوع العـام لمـصدر المعلـومات ، ويـؤخذ مـن قائمـة اسـتنادية عـادة مـا تـرد ملحقـة بالقـواعد الأنجلـو أمـريكية للفهرسـة نفسها . والـتحديدات العامـة للمـادة هـي عـبارة عـن كلمـات معبـرة مـثل: "نـص text " ، و" فـيلم متحـرك motion picture" ، و" تسجيل صـوتي sound recoding".

أمـا العـنوان المـوازي فهـو العـنوان نفسه إلا أنـه مكتوب بلغة أو هجائية أخـرى . والبـيانات الأخـرى للعـنوان ، غالـبًا مـا تكون بمثابة عنوان فرعـي . أمـا بـيان المـسؤولية فهـو تـسجيل للكـيان المـسؤول[١] عـن محـتوى العمـل الموصـوف . وهـناك قـواعد تـحدد مـن يمكـن اعتباره مسؤولاً عن المحتوى، وكيفية تحديـده فـي حالـة تعـدد المـسؤولين الموجـودين على العمل ، والكيفية التي تسجل بها أسماؤهم .

ومـن الملاحـظ أن عناصـر البـيانات فـي أي حقـل تفصـل بيـنها علامات ترقيم معتمدة على التقنين الدولي للوصـف الببليوجرافي. حيث يفصل العنوان مثلاً عن البيانات الأخرى للعنوان بالشارحة، كما يفصل عن بيان المسؤولية بشرطة مائلة. ويوضـح المثال التالي نموذج لحقل العنوان وبيان مسؤولية :

The book on the bookshelf/ by Henry Petroski.

وهناك مثال آخر أكثر تعقيداً يبدو على النحو التالي:

---

(٢) بالطبـع المسؤول مسؤولية فكـرية وليس مـاديـة حتى يتـم التمييـز بينـه وبين النـاشـر أو الطـابـع [المترجم] .

Proceedings of the Bicentennial Conference on Bibliographic control for the New Millennium: confronting the challenges of networked resources and the web Washington, D.C., November ١٥-١٧،٢٠٠٠/sponsored by the Library of Congress Cataloging Directorate; edited by Ann M. Sandberg-Fox.

ومن الملاحظ كذلك أن كل عنصر من عناصر البيانات التي يتكون منها حقل العنوان وبيان المسؤولية تُرمّز في حقول فرعية منفصلة ، وذلك داخل إطار الحقل " ٢٤٥ " في صيغة مارك٢١، كما تتطلب صيغة مارك نفسها من المفهرسين تقديم معلومات إضافية عن المحتوى. ويستخدم المؤشران الخاصان بالحقل ٢٤٥ لتحديد إذا ما كان يجب عمل مدخل إضافي للعنوان أم لا ، وما عدد الحروف التي سيتم تجاهلها عندما يتم الفرز حسب العنوان. ومن ثم سيظهر المثال الأول السابق في صيغة مارك٢١ على النحو التالي:

٢٤٥ ١٤ $a The book on the bookshelf / $c by Henry Petroski

ويتضح أن المؤشر الأول "١ " يشير إلى أن ثمة مدخلاً إضافياً بالعنوان[1]، في حين يشير المؤشر الثاني "٤" إلى أن عملية الترتيب يجب أن تبدأ بكلمة "book " أما أداة التعريف - The - والمسافة الفاصلة بينها وبين الكلمة الأولى - bookshlf - فيجب تجاهلهما[2]. أما العنوان نفسه فيظهر في

(١) يستعمل هذا المؤشر ليشير إلى أن لهذا العمل مدخلاً رئيسياً ، وهو في هذا المثال المؤهل، ومن ثم يتم إعداد مدخل إضافي بالعنوان [المترجم] .

(١) عندما يذكر المؤشر هكذا " ٤" يعني تجاهل أربع تمثيلات، ثم يبدأ الفرز بعد ذلك [المترجم].

الحقل الفرعي المسبوق بحرف "a" ، كما أن بيان المسؤولية يظهر في الحقل الفرعي المسبوق بحرف "c".

وعلى هذا النحو تعامل الحقول السبعة الأخرى. وتحدد القواعد الأنجلو أمريكية للفهرسة في طبعتها المنقحة عناصر البيانات التي يجب أن تتضمنها حقول الوصف، ومصادر المعلومات التي ينبغي أن تؤخذ منها بيانات كل حقل ، وقواعد كتابتها (مثل الاختصار والأحرف الكبيرة وقوائم الاستناد للمصطلحات) ، وعلامات الترقيم التي يجب استخدامها. أما مارك٢١ فيحدد طريقة تتويج الحقول ، والمؤشرات ، وتقنينات الحقول الفرعية[١] subfields code المستخدمة لترميز المحتوى. وباختصار فإن حقل الطبعة (الحقل رقم ٢٥٠ في مارك٢١) يمكن أن تحتوي على عناصر بيانات متصلة ببيانات الطبعة وببيانات المسؤولية المرتبطة بطبعات معينة.

وأما الحقل الثالث الخاص بالتفاصيل المحددة للمادة (أو نوع المطبوع) (الحقول : ٢٥٤ ، و ٢٥٥ ، و ٢٥٦ ، و ٣٦٢ في مارك٢١) فهو من الحقول غير المستخدمة مع بعض أشكال المواد، كما تستخدم تسميات محددة لبعض الأشكال الأخرى من المواد؛ حيث يسمى الحقل الثالث بالنسبة للمواد

(٢) يقصد بتقنينات الحقول الفرعية الرموز التي تستخدم لتحديد ذاتية الحقول الفرعية أو عناصر البيانات التي تتكون منها الحقول المتغيرة ، كما هو الحال مع حقل النشر ؛ حيث يتكون من ثلاثة عناصر أساسية : مكان النشر ، واسم الناشر ، وتاريخ النشر ، في هذه الحالة يسبق كل عنصر بتقنينة الحقل لتحديد هوية العنصر ، هكذا : **a** $ الرياض : **b**$ دار المريخ ، **c**$ ٢٠٠٥م [المترجم] .

الخرائطية - مثلاً- "حقل البيانات الرياضية" (الحقل رقم ٢٥٥ في مارك٢١) ، ويستخدم لتسجيل كل من: مقياس الرسم ، والإسقاطات ، والإحداثيات.

أما الحقل الرابع (الحقل ٢٦٠ في مارك٢١) فيستخدم للمعلومات المتعلقة بالنشر والتوزيع ، أحدهما أو كليهما ، ويشمل اسم الناشر أو الموزع، ومكان النشر أو التوزيع ، وتاريخ النشر والتوزيع وحقوق الطبع.

أما حقل الوصف المادي (الحقل ٣٠٠ في مارك٢١) فيتضمن قواعد تتفاوت تفاوتاً كبيراً حسب شكل المادة الموصوفة. وتشمل عناصر البيانات التي يمكن تسجيلها في هذا الحقل كلاً من: عدد الأجزاء ، والتفاصيل المادية الأخرى (مثل سرعة تشغيل التسجيلات الصوتية) ، وأبعاد المادة ، ووصف المواد المصاحبة.

أما حقل السلسلة (الحقول ٤٤٠ ، ٨٠٠، ٨١٠ ، ٨١١ ، ٨٣٠ في مارك٢١) فيستخدم في حالة ما إذا كانت المادة الموصوفة صادرة كجزء من سلسلة.

أما حقل التبصرات[1] (الحقول التي تبدأ برقم "٥" في مارك٢١) فيستخدم للمعلومات الوصفية المهمة التي لم ترد ضمن الحقول الأخرى. وتحدد القواعد لكل نوع من المواد التبصرات الخاصة بها ، وترتيب ظهورها، وفي بعض الحالات تحدد الصيغة الفعلية التي تكتب بها

(١) يسمى في بعض الترجمات بحقل الملاحظات [المترجم].

التبصرة[١]. وهذه التبصرات تسجل فيها معلومات مثل: محتوى العمل، والجمهور المستهدف، والأشكال الأخرى المتوافرة من العمل، وضوابط استخدامه.

أما الحقل الأخير (الحقل ٠٢٠ ، والحقل ٠٢٢ ، وحقول أخرى في مارك ٢١) فيستخدم لتسجيل الأرقام المعيارية مثل: الرقم المعياري الدولي للكتب ISBN ، والرقم المعياري الدولي للدوريات ISSN.

ويعنى الجزء الثاني من القواعد الأنجلو أمريكية للفهرسة في طبعتها المنقحة باختيار نقاط الإتاحة وصياغتها ، أو بالمداخل التي يمكن عن طريقها استرجاع تسجيلات الفهرسة ، ومع أنه يمكن في بيئة الحاسبات الحالية استخدام أي كلمة في أي عنصر من عناصر التسجيلة الببليوجرافية لأغراض الاسترجاع، حيث تعامل العناصر كلها كنقاط إتاحة في قواعد الفهرسة ، لكنها عادة ما تكون قابلة للبحث بطرق معينة ، كما هو الحال عند استخدام كشافات التصفح browse indexes. وهناك نوعان من نقاط الإتاحة : نقاط الإتاحة الببليوجرافية، ونقاط الإتاحة الموضوعية.

وتشمل نقاط الإتاحة الببليوجرافية أسماء المؤلفين وغيرها من الأطراف الأخرى المرتبطة بالعمل ، كما تشمل أسماء الهيئات المرتبطة بشكل ما بالعمل ، وكذلك أسماء السلاسل، والعناوين. ومن أهم المفاهيم المتداولة في هذا السياق مفهوم "المدخل الرئيسي " أو نقطة الإتاحة الرئيسية لتسجيلة الفهرسة ، والتي ترمز هكذا " ١XX " في صيغة مارك٢١؛ أما جميع

(٢) تعرف في اصطلاح المفهرسين بالتبصرة المقننة على عكس التبصرة الحرة [المترجم].

نقاط الإتاحة الببليوجرافية الأخرى فهي تعد مداخل إضافية ، وترمز في الحقول: ٤xx ، أو ٧xx ، أو ٨xx. (مع ملاحظة أن الترميز " nXX " يعني أي تاج حقل يبدأ بالخانة " n " . فمثلاً حقل " ١xx " ، قد يكون : ١٠٠ ، أو ١١٠ ، أو ١٢٠ ، أو ١٣٠).

كذلك تشتمل القواعد الأنجلو أمريكية للفهرسة على قواعد دقيقة لتحديد المدخل الرئيسي المعتمد على مفهوم التأليف authorship - ذلك المفهوم الغامض نسبياً - وليس على ما هو أكثر وضوحاً كالعنوان، ومع الأهمية الكبيرة التي كان يكتسبها المدخل الرئيسي في نظم الفهارس البطاقية، فإن الحاجة إليه في البيئة الحالية لم تعد بتلك الأهمية ؛ حيث يرى البعض أنه لا تزال هناك فائدة لنقطة الإتاحة الأولى الأساسية لتحقيق بعض الوظائف المحددة مثل : عرض التسجيلات الببليوجرافية citations ، وترتيب المواد المسترجعة، في حين يحتج آخرون بأن تعقيدات تحديد هذه النقاط تقلل من فائدتها.

وعلى أية حال ، ومع أنه يمكن التفرقة بين كل من المداخل الرئيسية والمداخل الإضافية؛ إلا أن قواعد صياغتها واحدة. وبالرغم من أن القواعد الخاصة برؤوس أسماء الأشخاص ، والأسماء الجغرافية ، وأسماء الهيئات ، وأحد الأنواع الخاصة من العناوين والذي يعرف "بالعناوين المقننة uniform title " قد باتت راسخة ، فإن هناك بالنسبة لأسماء الأشخاص قضيتين : قضية اختيار الاسم ، وقضية شكل الاسم. وبالنسبة لاختيار الاسم فإن القاعدة العامة هي اختيار الاسم الذي عادة ما يعرف به الشخص (مثلاً

جيمي كارتر Jimmy Carter بدلاً من جيمس إيرل كارتر Jemes Earl Carter) ، ولكن هناك قواعد إضافية لمعالجة الحالات الغامضة نوعاً ما ، مثل الحالات التي يشار فيها للمؤلف باسم أو أسماء مستعارة ، أو الحالات التي غير فيها المؤلفون أسماءهم. وبمجرد اختيار الاسم نفسه، فإن شكله ينبغي أن يحدد ، مع مراعاة بعض الاعتبارات مثل : الاكتمال ، واللغة ، وترتيب المدخل بالنسبة للأسماء المستخدمة كنقاط إتاحة ، مع اللجوء لإضافة تاريخ الميلاد أو الوفاة - أحدهما أو كليهما - لتمييز الأسماء المتشابهة بعضها عن بعض.

ولإكمال التسجيلة الببليوجرافية هناك مستويان آخران من مستويات وصف المصادر لم تتناولهما القواعد الأنجلو أمريكية للفهرسة في طبعتها المنقحة ، ومن الضروري عدم تجاهلهما ، وهما: الفهرسة الموضوعية والتصنيف. حيث تعنى الفهرسة الموضوعية بتعيين نقاط إتاحة موضوعية ؛ فمعظم التسجيلات الببليوجرافية الخاصة بالأعمال العلمية غير الأدبية nonfiction تشتمل على رأس موضوع واحد على الأقل مأخوذة من لغة مقيدة controlled vocabulary . وفي هذا الشأن تستخدم المكتبات الكبرى والأكاديمية والبحثية قائمة رؤوس موضوعات مكتبة الكونجرس LCSH بشكل كبير، بينما تلجأ المكتبات الصغرى ومكتبات المدارس إلى استخدام قائمة سيرز sears لرؤوس الموضوعات . وقد تستخدم أنواع متخصصة من المكتبات لغات أخرى مقيدة تتفق ومجالها الموضوعي وعدد المستفيدين منها ،مثال ذلك قائمة رؤوس الموضوعات الطبية MeSH التي أعدتها المكتبة الطبية الوطنية ، أو مكنز جيتي للفنون و العمارة . وتجدر الإشارة

إلى أن المصطلحات الموضوعية ترمز كتيجان ٦XX في صيغة مارك ، مع تمييز كل منها بنوع المصطلح الموضوعي (مثال ذلك الاسم الشخصي، أو المصطلح الموضوعي ، أو المنطقة الجغرافية)، واللغة المقننة المستخدمة .

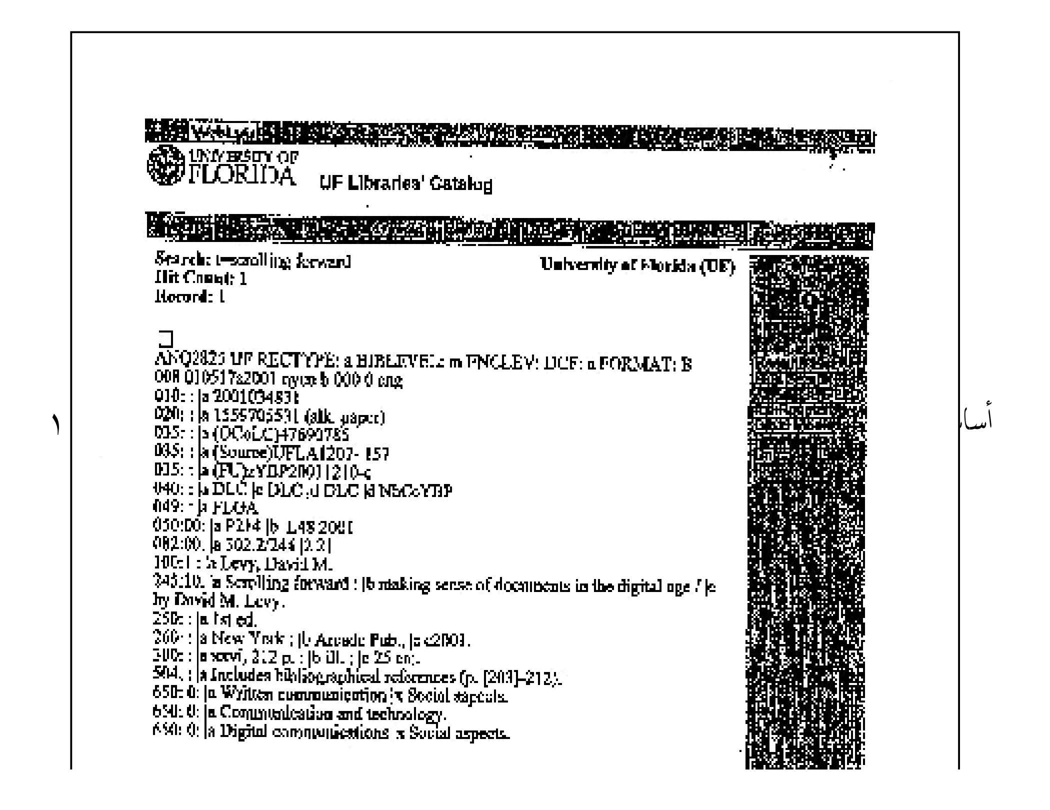
أما التصنيف أو استخدام الرموز اعتمادًا على خطة تصنيف مختارة فيخدم عدة أغراض ، فكون رقم التصنيف جزءاً من رقم الطلب أو الوسيمة المستخدمة لتحديد مكان مصدر المعلومات على الرف، فإنه يعمل على تجميع المواد المتشابهة موضوعيا معاً ، لذلك فإن قائمة الرف - المطبوعة أو الإلكترونية - والمرتبة وفقاً لترتيب جميع المواد الفعلية التي تقتنيها المكتبة على الرفوف، يستخدمها العديد من المكتبات أداة لضبط مقتنياتها AN INVENTORY CONTROL DEVICE ، كما يمكن أيضاً استخدام رقم التصنيف كأحد أساليب الوصول الموضوعي في عمليات البحث والاسترجاع. وكما هو الحال مع رؤوس الموضوعات ، ثمة تفاوت بين المكتبات في تفضيلها استخدام خطط التصنيف كأسس لأرقام الطلب الخاصة بها، ففي الوقت الذي تميل فيه المكتبات الكبرى وخاصة الأكاديمية منها لاستخدام خطة تصنيف مكتبة الكونجرس ، تفضل مكتبات أخرى استخدام تصنيف ديوي العشري في هذا الصدد ، بل تفضل المكتبات الأوربية عموماً استخدام التصنيف العشري العالمي .

ويوضح الشكل رقم (١/٦) شاشة لتسجيلة ببليوجرافية بصيغة مارك٢١. ونلاحظ فيها أن ما يقرب من نصف الحقول هي حقول الضبط أو حقول " OXX "؛ إذ تشتمل على أرقام ضبط ورموز معالجة متنوعة، حيث يشتمل الحقل ٠٥٠ على سبيل المثال على رقم تصنيف مكتبة الكونجرس، كما

يشتمل الحقل ٠٨٢ على رقم تصنيف ديوي العشري، ثم يخصص الجزء المتبقي من التسجيلة للبيانات الوصفية البسيطة الخاصة بالعناصر التالية : المداخل الرئيسية، وبيانات العنوان، وبيان الطبعة، والنشر، والوصف المادي، ثم تبصرة ببليوجرافية، تتبعها ثلاثة رؤوس موضوعات مأخوذة من قائمة رؤوس موضوعات مكتبة الكونجرس. أما الشكل رقم (٢/٦) فيعرض التسجيلة نفسها التي عرضت في الشكل السابق ، ولكن كما تظهر في شاشة فهرس عام مباشر لإحدى المكتبات .

فهرسة المصادر الإلكترونية :

لكي تستوعب كل من صيغة مارك وقواعد الفهرسة وصف المصادر الإلكترونية ، أجريت عليهما تعديلات جوهرية ، فعندما نشرت القواعد الأنجلو أمريكية للفهرسة في أول مرة عام ١٩٧٨م لم تكن تشتمل إلا على فصل واحد عن ملفات البيانات المقروءة آلياً (الفصل ٩) ، وقد نقح هذا الفصل في عام ١٩٨٨م وتغير عنوانه ليصبح " ملفات الحاسب" ، كذلك وافقت اللجنة الاستشارية لمارك الأمريكي على إضافة صيغة ملفات الحاسب إلى صيغة مارك الأمريكي في عام ١٩٨١م.


UNIVERSITY OF FLORIDA UF Libraries' Catalog

Search: t=scrolling forward
Hit Count: 1
Record: 1

University of Florida (UF)

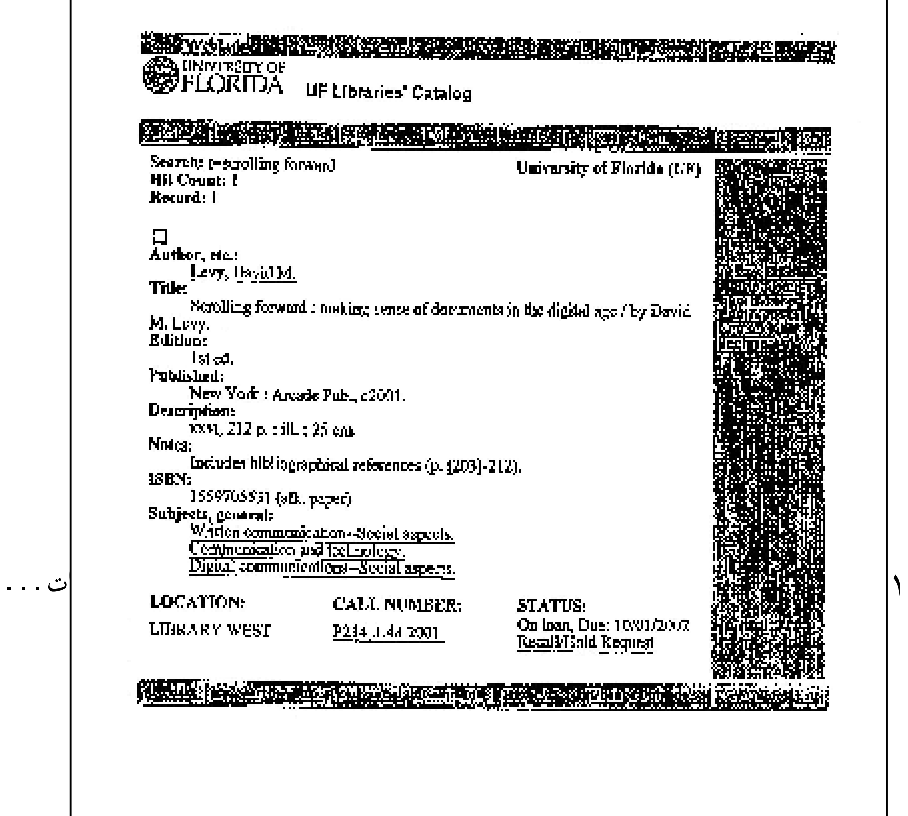
ANQ2825 UF RECTYPE: a BIBLEVEL: m ENCLEV: DCF: a FORMAT: B
008 010517s2001 nyua b 000 0 eng
010: : |a 2001034831
020: : |a 1559705531 (alk. paper)
035: : |a (OCoLC)47690785
035: : |a (Source)UFLA1207- 157
035: : |a (FU)zYBP2001 1210-c
040: : |a DLC |c DLC |d DLC |d NScCoYBP
049: : |a FLGA
050:00: |a P214 |b .L48 2001
082:00: |a 302.2/244 |2 21
100:1 : |a Levy, David M.
245:10: |a Scrolling forward : |b making sense of documents in the digital age / |c by David M. Levy.
250: : |a 1st ed.
260: : |a New York : |b Arcade Pub., |c c2001.
300: : |a xxvi, 212 p. : |b ill. ; |c 25 cm.
504: : |a Includes bibliographical references (p. [203]-212).
650: 0: |a Written communication |x Social aspects.
650: 0: |a Communication and technology.
650: 0: |a Digital communications |x Social aspects.

شكل رقم ١/٦

شاشة لتسجيلة ببليوجرافية في صيغة مارك ، من مكتبة جامعة فلوريدا

الشكل رقم (١/٦)

شاشة لتسجيلة ببليوجرافية في صيغة مارك، من مكتبة جامعة فلوريدا


UNIVERSITY OF FLORIDA UF Libraries' Catalog

Search: (=scrolling forward) University of Florida (UF)
Hit Count: 1
Record: 1

Author, etc.:
Levy, David M.
Title:
Scrolling forward : making sense of documents in the digital age / by David M. Levy.
Edition:
1st ed.
Published:
New York : Arcade Pub., c2001.
Description:
xxvi, 212 p. : ill. ; 25 cm.
Notes:
Includes bibliographical references (p. [203]-212).
ISBN:
1559705531 (alk. paper)
Subjects, general:
Written communication--Social aspects.
Communication and technology.
Digital communications--Social aspects.

LOCATION:	CALL NUMBER:	STATUS:
LIBRARY WEST	P214 .L44 2001	On loan, Due: 10/01/2007 Recall/Hold Request

ت . . .

الشكل رقم (٢/٦)

شاشة للتسجيلة الببليوجرافية المعروضة في الشكل السابق رقم (١/٦) من الفهرس العام المتاح للجمهور للمكتبة نفسها

وكما يشير العنوان "ملفات الحاسب " ـ فقد ركز مارك وقواعد الفهرسة بداية على ملفات البيانات (كبيانات الدراسات المسحية في العلوم الاجتماعية)، والمطبوعات المنشورة في أشكال مادية (كالبرمجيات المحملة على الأقراص المرنة، والموسوعات المحملة على الأقراص المدمجة). وإلى حد بعيد يمكن النظر إلى التحولات التاريخية للفهرسة في تسعينات القرن العشرين على أنها محاولة لمواكبة النمو المتسارع الذي تشهده المصادر الرقمية المتاحة عن بعد ، والتي كانت تمثلها في بداية الأمر الملفات والخدمات التي يتم الوصول إليها عن طريق بروتوكولات الإنترنت مثل : بروتوكول نقل الملفات FTP ، و تلنت TELNET ، ثم تمثلت لاحقاً في جميع أنواع المصادر المتاحة على الوب.

ويعد مشروع (أو سي إل سي لمصادر الإنترنت OCLC Internet Resources Project) الذي مولت تنفيذه وزارة التعليم الأمريكية وجرى تنفيذه خلال الفترة من ١٩٩١ إلى ١٩٩٣ بقيادة مارتن ديلون Martin Dillon ، من التطورات البارزة في هذا الشأن. وقد ركزت المرحلة الأولى من هذا المشروع على تصنيف أنواع المصادر المتاحة على الإنترنت وقتئذ، هذا

في الوقت الذي اهتمت فيه المرحلة الثانية بتجربة تطبيق القواعد الأنجلو أمريكية للفهرسة وصيغة مارك الموجودة في ذلك الوقت على فهرسة عينة من تلك المصادر. وقد أثمرت التجربة عن حصر المشكلات التي تتصل بحقول الوصف سواء في صيغة مارك الأمريكي أو القواعد الأنجلو أمريكية للفهرسة، كما كان من الثمرات المباشرة للمشروع ذلك الدليل الإرشادي لنانسي أولسون المسمى " Olson's Cataloging Internet Resources; A manual and Practical Guide"، والذي كان له أثره الواضح في تشجيع فهرسة مصادر الإنترنت ،بل في تشكيل توجهات الممارسات المبكرة لفهرسة هذا النوع من المصادر(١)، كما كان من النتائج الأخرى للمشروع إضافة حقل " الموضع الإلكتروني وكيفية الإتاحة electronic location and access " (الحقل ٨٥٦) إلى صيغة مارك الأمريكي.

وإذا كان هذا الحقل (رقم ٨٥٦) يستخدم الآن لتسجيل المحددات الموحدة لمواقع المصادر URLs الخاصة بالمصادر المتاحة على الوب ، فإنه من ناحية قد ظهر قبل اتساع شبكة الوب ؛ حيث كان قد اقترح في عام ١٩٩٢م وأقر في عام ١٩٩٣م - أي في العام نفسه الذي أطلق فيه المتصفح السابق المعروف بِـ"موزايك"- وقد كان هذا الحقل كذلك قد استحدث في الأساس لتسجيل سبل الإتاحة أو الوصول التي كانت سائدة عند دراسة مشروع مصادر الإنترنت وهي : نقل الملفات ، والبريد الإلكتروني ،

---

(١) Nancy B. 'Olson, ed., Cataloging Internet Resources: A Manual and Practical Guide, ٢nd ed., available t http://www.purl.org/oclc/cataloging-internet. Accessed ١٨ June ٢٠٠٢.

وتسجيل الدخول عن بعد. أما الحقل الفرعي الذي يرمز له بالحرف u والخاص بالمعرف الموحد لموقع المصدر URL فقد أضيف كتعديل للحقل ٨٥٦ بعد إضافته إلى الصيغة بستة أشهر .

وهناك تغييرات أخرى استحدثت على مارك في بداية التسعينات من القرن العشرين، منها إضافة مجموعة من القيم المرمزة للوصف المادي لملفات الحاسب (الحقل ٠٠٧) ، وإضافة رمز" لحالة الإتاحة عن بعد remote "، وإضافة حقول بيانات لوصف النظم والخدمات المباشرة.

أما التغييرات التي استحدثت على تقنينات الفهرسة فقد تأخرت نظراً للمعالجة المعقدة التي تحكم القواعد الأنجلو أمريكية للفهرسة من ناحية، وللطبيعة المتحفظة لرواد الفهرسة من ناحية أخرى. ففي عام ١٩٩٧م صدر التقنين الدولي للوصف الببليوجرافي (المصادر الإلكترونية) ISBD (ER) ، والذي اتخذ أساساً للمراجعة الكاملة للفصل التاسع من القواعد الأنجلو أمريكية الصادر في ٢٠٠١م مع تغيير عنوانه من "ملفات الحاسب " إلى "المصادر الإلكترونية" ، واتسع مجال تغطيته ليشمل المصادر المتاحة بشكل مباشر(١) direct access resources (المصادر ذات وسائط التحميل المادية التي يتم إدخالها في الحاسب مثل الأقراص المدمجة وأشرطة التخزين) ، والمصادر المتاحة عن بعد remote access resources

(١) يقصد بالمصادر التي تتاح على الخط المباشر تلك المصادر التي يمكن اختزانها على وسيط مادي إلكتروني ويمكن إتاحتها عبر الشبكات المحلية وليس عن طريق الإنترنت [المترجم] .

(المصادر التي ليس لها وسائط تحميل مادية متصلة بالشبكة أو من خلال وسائط تخزين مرفقة بشكل مباشر).

ويؤخذ عنوان المصدر الإلكتروني (الحقل ٢٤٥ من مارك٢١) من المصدر الرئيسي للمعلومات، الذي هو عبارة عن أي " علامة أو شاهد يرد في بداية العمل وبشكل رسمي " في المصدر نفسه ، ويشمل ذلك: شاشة العنوان ، أو الصفحة الرئيسية ، أو القائمة الرئيسية، أو المعلومات التي تظهر في بداية العرض للمصدر، أو ترويسة الملف ، أو الحامل المادي، أو ما وراء البيانات المرمزة بتيجان الميتا <meta> في الصفحات المهيأة بصيغة لغة ترميز النصوص الفائقة html ، أو لغة الترميز الموسعة xml . وفي حالة وجود أكثر من مصدر من هذه المصادر ينبغي استخدام المصدر "الأكمل" أو الأكثر اكتمالاً ، مع ضرورة تسجيل المصدر الذي أخذ منه العنوان في تبصرة توضح ذلك. ويعد مصطلح "مصدر إلكتروني electronic resource " هو التحديد العام للمادة GMD المعبر عن جميع أنواع المصادر الإلكترونية .

وإذا كان المصدر الإلكتروني يحتوي على بيانات تتصل بالطبعة كما يظهر ذلك واضحاً من كلمات مثل: "طبعة" ، أو "إصدارة " ، أو "نسخة" ، أو "تحديث"، ينبغي تسجيل هذه المعلومات في حقل الطبعة (الحقل ٢٥٠ في مارك٢١) ، هكذا : " version ٥,٢٠ ".

ويسمى الحقل الثالث بالنسبة للمصادر الإلكترونية (الحقل ٢٥٦ في مارك ٢١) "نوع المصدر ومداه type and extent of resource area"، وتخصص

بموجبه ثلاثة مصطلحات فقط لتعريف نوع المصدر هي : "بيانات إلكترونية electronic data "، و"برنامج إلكتروني (أو برامج إلكترونية) electronic program(s) "، و "بيانات وبرنامج (أو برامج) إلكترونية electronic data and program(s) "، وهذا مما يؤسف له لأن ذلك يسبب إرباكاً للمستفيد، وخاصة عندما توصف مصادر أخرى كالدوريات الإلكترونية ومواقع الوب بمصطلح "بيانات إلكترونية" ؛ لذلك فقد رأت مكتبة الكونجرس في تطبيقها لذلك أن تحذف مثل هذه المعلومات غير المعبرة، كما يحدد التقنين الدولي للوصف الببليوجرافي (المصادر الإلكترونية) ISBD (ER) قائمة أطول وأكثر واقعية تشتمل على مصطلحات أخرى مثل : "دورية إلكترونية"، و"بيانات إلكترونية مصورة" ، و "خدمة إلكترونية مباشرة". ولعل بإمكان الإصدارات المستقبلية من تقنينات الفهرسة أن تضم إليها هذه القائمة المطولة، أو - وهذا هو الأنسب - أن يلغى الحقل الثالث الخاص بالمصادر الإلكترونية.

كذلك يتم إعطاء بيانات النشر والتوزيع (الحقل ٢٦٠ في مارك٢١) للمصادر الإلكترونية بالطريقة نفسها التي تعطى بها للأنواع الأخرى من المصادر؛ فجميع المصادر المتاحة عن بعد حتى الصفحات الرئيسية الشخصية تعامل على أنها مصادر منشورة .

ويسمح بالوصف المادي (الحقل ٣٠٠ في مارك ٢١) فقط للمصادر المتاحة بشكل مباشر ، وليس للمصادر التي تتاح عن بعد ، وتشمل التبصرات المسموح بها طبيعة الملف وحجمه (حيث يمكن وصف نوع المصدر بشكل رمزي أقل مما هو عليه في الحقل الثالث) ، بالإضافة إلى أية متطلبات (مادية أو

برمجية) لاستخدام هذا الملف. ويتطلب الأمر وجود تبصرة تحدد طريقة الإتاحة أو الوصول للملف "mode of access"، مثال ذلك "الشبكة العنكبوتية " وذلك بالنسبة للمصادر التي يمكن الوصول إليها عن بعد، وكذلك يتطلب الأمر إعطاء تبصرة تبين تاريخ فهرسة المصدر.

وتشترك المصادر الإلكترونية بالطبع مع الأشكال الأخرى من مصادر المعلومات في خصائص معينة ، فقد تكون هذه المصادر الإلكترونية مواد خرائطية أو مخطوطات أو تسجيلات صوتية ، أو مواد نصية أو غير نصية. ولذلك تدعو قواعد الفهرسة إلى معالجة جميع الأشكال التي تمثلها هذه المصادر، بحيث يمكن - مثلاً - أن توصف الدوريات الإلكترونية وفقاً لقواعد وصف المسلسلات من ناحية، وقواعد وصف المصادر الإلكترونية من ناحية أخرى، كما أن هناك بعض الأشكال من المصادر الرقمية شائعة الاستخدام مثل مواقع الوب وقواعد البيانات المتاحة على الخط المباشر ، لا تتوافر لها نظائر غير إلكترونية، ولم تعالجها قواعد الفهرسة الحالية المعالجة المناسبة ، فمن الملاحظ أن هذه المصادر ليست ذات طبيعة ثابتة كما هو الحال مع المنفردات، كما أنها لم تصدر في شكل تتابعي كالمسلسلات ، غير أنها تحدث بشكل مستمر، ومن ثم فإنها تتشابه في كثير من الجوانب مع المطبوعات التي تصدر في أوراق سائبة، لهذا يحاول المفهرسون الاستئناس بالقواعد الخاصة بتلك المطبوعات ذات الأوراق الحرة عند وصف مثل هذه المصادر.

وتتمثل آخر تحديثات الطبعة المنقحة من القواعد الأنجلو أمريكية للفهرسة في الفصل الثاني عشر الخاص " بالمسلسلات " (ALA, ٢٠٠٢) ، وقد حرصت مراجعة هذا الفصل على التفرقة بين نوعين من "المصادر المستمرة continuing resources"

، همـا : المـصادر المتعاقـبة successive (أي الـصادرة فـي أجـزاء منفـصلة كالمسلـسلات والـسلاسل) ، والمـصادر المـتكاملة integrating (أي التـي تحـدث باستمرار كالمطبوعات ذات الأوراق الحرة ومواقع الوب)، ولعل في ذلك ما يوضح الصورة ويبسطها عند فهرسة تلك المصادر الهامة.

وهناك قضية هامة أخرى رئيسية من قضايا الفهرسة تثيرها المصادر الرقمية ، وهي قضية كيفية التعامل مع الإصدارات المختلفة للعمل الواحد ؛ فمن الشائع أن تأتي الكيانات الرقمية وكأنها إصدارات إلكترونية لكيانات مادية (كالمواد الوثائقية والمواد الخرائطية وغيرهما) ، يضاف إلى ذلك أن محتوى هذه الكيانات نفسها قد يصدر في عدة صيغ رقمية مختلفة (كصيغة النص الفائق HTML ، وصيغة الوثيقة المحمولة أو ما تعرف بالبي دي إف PDF ، وصيغة محرر النصوص أو ما يعرف بميكروسوفت وورد MS Word) ، وهذا الوضع هو ما يضفي قدراً كبيراً من الحساسية على كيفية التعامل مع تلك الإصدارات ، وخاصة إذا ما وضع في الاعتبار محاولة التخفيف من العبء الملقى على عاتق المفهرس ،وفي الوقت نفسه الارتقاء بمعدلات شفافية الفهرس للمستفيد .

وثمة تحول في الرؤى تجاه هذه القضية في ظل ما أثير في سياق المتطلبات الوظيفية للتسجيلة الببليوجرافية(١) FRBR من ضرورة التمييز بين كل من الأعمال works ، وأساليب التعبيرات expressions ، ووسائط العرض manifestations ، والمـواد أو الـوحدات items ، وكيفـية تعامـل تقنيـنات الفهرسة ومارك مع هذه الجوانب من جوانب التفرقة في المستقبل.

(١) تلك التي أقرها الاتحاد الدولي لجمعيات المكتبات ومعاهدها [المترجم].

وحسبنا القول إن من أكثر المقولات تداولاً في هذا الصدد أن الشيء الثابت الوحيد في البيئة الرقمية اليوم هو التغير ، ومن المؤكد أن الأشكال التي تتوافر فيها المعلومات الرقمية الآن، والسبل التي يسلكها المستفيدون للوصول إلى تلك المصادر سوف تتغير ، كما أن فهرسة المكتبات التقليدية سوف تتغير كذلك حتى يمكنها استيعاب مستجدات العصر، كما ينبغي على المهتمين بذلك أن يتابعوا أنشطة الجهات الرئيسية التي تتبنى تلك المستجدات في هذا الصدد، فقد تبنت مكتبة الكونجرس مارك ٢١ بإشراف لجنة مارك الاستشارية[١]، كما أن متابعة التحديثات لقواعد الفهرسة هو ما انصبت عليه جهود لجنة التوجيه المشتركة لتنقيح القواعد الأنجلو أمريكية للفهرسة، وهي عبارة عن مجموعة دولية قناتها الرئيسية لإدخال تلك التعديلات هي لجنة الفهرسة التابعة لجمعية المكتبات الأمريكية المعنية بقواعد الوصف ونقاط الإتاحة ALA Committee on Cataloging: Description and Access ، والتي تعرف اختصاراً ب (CC:DA)[٢].

READINGS :

---

(١) See the Link to "MARC Development" from the Library of Congress, Network Development and MARC Standards Office, MARC Standards page at http://lcweb. loc.gov/marc/

(٢) Joint Steering Committee for Revision of Anglo-American Cataloguing Rules (home page), available at http://www.nlc-bnc.ca/jsc/. Accessed ٢٦ August ٢٠٠٢; Committee on Cataloging: Description and Access (home page), available at http://www.ala.org/alcts/ organization/ccs/ccda/ccda.html. Accessed ٢٦ August ٢٠٠٢.

- Anglo-American Cataloguing Rules. ٢nd ed. Prepared by the American Library Association, the British Library, the Canadian Committee on Cataloguing, the Library Association, the Library of Congress. Edited by Michael Gorman and Paul W. Winkler. Chicago: American Library Association, ١٩٧٨.

  Although you may not want to read this cover to cover, a careful look at the organization and content of AACR٢ will go a long way toward conveying the nature and spirit of library cataloging rules.

- International Federation of Library Associations and Institutions. ISBD(ER): International Standard Bibliographic Description for Electronic Resources. Available at http://www.ifla.org/VII/sI٣/pubs/isbd.htm.

- Olson, Nancy B., ed. Cataloging Internet Resources: A Manual and Practical Guide. ٢nd ed. Dublin ,Ohio: OCLC, ١٩٩٧. Available at http://www.purl.org/oclc/cataloging-internet.

  How to apply AACR٢ and ISBD(ER) to Internet-accessible resources, written before the ٢٠٠١ revision of chapter ٩.

- Svenonius Elaine. The Intellectual Foundation of Information Organization. Cambridge, Mass . MIT press, ٢٠٠٠.

A difficult but illuminating work well on its way to becoming a classic, this book gives a conceptual framework for the process of organizing information in the tradition of Anglo-American descriptive and subject cataloging.

الفصل السابع

ترويسة مبادرة ترميز النص

مبادرة ترميز النص Text Encoding Initiative أحد المشروعات الدولية التي دشنت في عام ١٩٨٧م برعاية مشتركة من كل من جمعية الحاسبات في مجال الإنسانيات ، وجمعية اللغويات الحاسوبية ، وجمعية الحوسبة الأدبية واللغوية ، بهدف تطوير إرشادات لتطويع لغة الترميز المعيارية العامة SGML لترميز النصوص الإلكترونية، وتشجيع استخدامها في تبادل الدراسات العلمية المتخصصة في مجال الإنسانيات. وكان من ثمرات هذه المبادرة إصدار "إرشادات ترميز النصوص الإلكترونية وتبادلها Guidelines for Electronic Text and Interchange " التي نشرت في ثلاث إصدارات رئيسية: إصدارة عام ١٩٩٩م، وتعرف بمبادرة ترميز النص - ب ٣ TEI P٣ ، وتركز على التعريف بكيفية ترميز النصوص وفقاً للغة الترميز المعيارية العامة، وإصدارة عام ٢٠٠٢م الحالية التي حلت محل الإصدارة الأولى، وتعرف بمبادرة ترميز النص - ب ٤ TEI P٤ ، وهي طبعة منقحة بالكامل من الإصدارة الأولى لتدعم صيغة لغة الترميز الموسعة XML ، مع استمرار توافقها مع مبادرة ترميز النص - ب ٣ TEI P٣، وهناك الإصدارة الثالثة التي تعرف بمبادرة ترميز النص المبسطة TEI Lite وهي الأبسط والأسهل استخداماً مقارنة بالإصدارتين السابقتين. وتجدر الإشارة إلى أن ما سنطرح من أمثلة ومناقشات في الفقرات التالية سيعتمد على الإصدارة الأولى من تلك المبادرة (أي مبادرة ترميز النص - ب ٣ TEI P٣).

جاءت " إرشادات مبادرة ترميز النص TEI Guidelines " عندما نشرت لأول مرة عام ١٩٩٤م في ١٣٠٠ صفحة، وكانت تزن سبعة

أرطـال(١)، ويعود كبر حجمها هذا إلى أنها تحاول التعامل مع ترميز كثير من الأنواع المختلفة لنصوص التحليلات التاريخية والأدبية واللغوية ، مع ملاحظة أن الجزء الوحيد الذي يختص بما وراء البيانات من هذه الإرشادات هو الفصل الخامس منها ، ويعنى بتعريف ترويسة مبادرة ترميز النص TEI header . وقد أشير إلى طبيعة هذا الفصل وأهدافه في سياق الإصدارة الثانية من المبادرة (مبادرة ترميز النص - ب ٤ TEI P٤) على النحو التالي:

يعالج هذا الفصل قضايا وصف العمل المرمز ، بحيث يوثق نص العمل نفسه، ومصدره، وأسلوب ترميزه ، ومراجعاته توثيقاً وافياً. ويكتسب هذا التوثيق القدر نفسه من الأهمية لكل من الباحثين الذين يستخدمون تلك الأعمال، والبرمجيات التي تعالجها، وكذلك المفهرسين في المكتبات ودور المحفوظات ، كما تسعى ما تقدم من توصيفات وتوضيحات إلى توفير نظير إلكتروني لصفحة عنوان العمل المطبوع ، كما أنها تعد كذلك أشبه بأدلة الترميز code books والأدلة التقديمية التي جرت العادة أن تصاحب مجموعات البيانات الإلكترونية(٢).

ويبدو واضحاً من سياق الاقتباس السابق أن واضعي مسودة إرشادات مبادرة ترميز النص - باعتبارهم باحثين في الإنسانيات - كانوا هم أنفسهم ممن يستخدمون المكتبات استخداماً مكثفاً، كما أنهم كانوا على دراية كبيرة

---

(١) الرطل الواحد = ٠,٤٥٣ كلغم [المترجم] .

(١) C. M. Sperberg-McQueen and L. Burnard, eds., TEl P٤: Guidelines for Electronic Text Encoding and Interchange, Text Encoding Initiative Consortium, XML Version (Oxford, Providence, Charlottesville, Bergen:].٠٠٢), available at http://www.teic.org/Guidelines٢/index.html. Accessed ١٩ June ٢٠٠٢.

باحتياجات تلك المكتبات. ومع بداية استخدام الحاسبات في مجال الإنسانيات قام بعض هؤلاء الباحثين (أو مساعدوهم من طلاب الدراسات العليا) بتحويل نصوص من هذا المجال إلى صيغ إلكترونية ، كما قاموا بترميزها بشكل يتفق وأهدافهم البحثية ، وعليه افترض واضعو مبادرة ترميز النص أن الباحثين في هذا المجال سينشئون نصوصاً مرمزة اعتماداً على تلك المبادرة بالإضافة إلى الترويسات الخاصة بهذه النصوص ، ومن ثم سيعتمد المكتبيون على تلك الترويسات عند فهرسة تلك النصوص التي ستعرض بدورها في فهارس المكتبات ؛ من هنا حرصوا على أن تصمم بنية الترويسة في شكل يتوافق مع التقنين الدولي للوصف الببليوجرافي ISBD ، ولكن مع عدم التقيد بحرفية القواعد الأنجلو أمريكية للفهرسة فيما يتصل بعرض محتوى تلك الترويسات.

تتكون ترويسة مبادرة ترميز النص من أربعة أقسام رئيسية ، لكن بحكم أنها تعد من قبيل معرف نوع المستند الخاص بلغة الترميز المعيارية العامة SGML DTD، فقد سميت تلك الأقسام في هذا السياق بمسميات تيجان تلك اللغة SGML tags ، وجاءت الأقسام الأربعة على النحو التالي:

- قسم[1] وصف الملف <fileDesc> ، ويشتمل على الوصف الببليوجرافي للنص الإلكتروني.

(١) تسمية الأقسام بهذه المسميات ، أي قسم كذا ..... من اجتهاد المترجم ، وما ورد في الأصل مجرد ذكر التيجان فقط ، هكذا : <fileDesc> ... إلخ ، وبدون ذكر كلمة (قسم) باللغة الإنجليزية [المترجم].

- قسم وصف الترميز <encodingDesc > ، ويصف العلاقة بين النص الإلكتروني والمصدر (أو المصادر) التي اشتق منها هذا النص.

- قسم وصف الخصائص غير الببليوجرافية <profileDesc > ، ويصف الجوانب غير الببليوجرافية للنص ، وخاصة اللغات المستخدمة ، والموقف الذي أعد فيه النص ، والواصفات الموضوعية والنوعية التي تعبر عن محتواه.

- قسم وصف المراجعة <revisionDesc > ويشتمل على تاريخ مراجعة النص .

ومع أن الترويسة ما هي إلا مجموعة فرعية من مواصفة مبادرة ترميز النص الكاملة ، فإنها وعلى ما قد تنطوي عليه من التعقيد، توفر عدداً من الخيارات ؛ حيث يلاحظ أن معظم عناصرها اختيارية، كما أن هناك أساليب بديلة لتمثيل المحتوى نفسه ، حيث يمكن - على سبيل المثال - تمثيل قيمة كل عنصر في جملة نصية بسيطة متصلة ، أو في جملة نصية تأتي في شكل فقرات ، أو في سلسلة من العناصر الفرعية المحددة للمحتوى بشكل أكثر تخصيصاً.

وصف الملف :

قسم وصف الملف <fileDesc> هو القسم الوحيد الإجباري في ترويسة مبادرة ترميز النص ، وقد صمم بشكل صريح ليتوافق وبنية التقنين الدولي للوصف الببليوجرافي ، ويشتمل هذا القسم على سبعة أقسام فرعية أو

حقول[١]، جاءت على النحو التالي:

- بيانات العنوان < titleStmt > وهو خاص بالمعلومات عن عنوان العمل وبيانات المسؤولية الفكرية .
- بيانات الطبعة < editionStmt > وهو خاص بالمعلومات عن هوية الطبعة التي صدر بها النص.
- بيانات الحجم < extentStmt > وهو خاص بالحجم التقريبي للعمل .
- بيانات النشر < publicationStmt > وهو خاص بالمعلومات عن نشر النص الإلكتروني وتوزيعه.
- بيانات السلسلة < seriesStmt > وهو خاص بالمعلومات عن السلسلة التي ينتمي إليها النص.
- بيانات التبصرات < notesStmt > وهو خاص بالتبصرات التي تقدم معلومات إضافية عن النص.
- وصف المصدر < sourceDesc > وهو عبارة عن الوصف البيليوجرافي للنص الأصلي (أو النصوص الأصلية) التي اشتق منها النص الإلكتروني.

ومن الواضح أن درجة التطابق بين الحقول الستة الأولى من هذا القسم

(١) ذكرت في الأصل أقسام فرعية subsections ، لكن رأيت أن أسميها حقولاً تمشياً مع مصطلحات الفهرسة، وخاصة أن هذا القسم قصد به أن يكون مطابقاً لبنية التقنين الدولي للوصف الببليوجرافي كما أشير إلى ذلك في متن النص [المترجم].

، وستة حقول من حقول الوصف في القواعد الأنجلو أمريكية للفهرسة المبنية على التقنين الدولي للوصف الببليوجرافي (انظر الفصل السادس من هذا الكتاب) لم تأت مصادفة أو عن غير قصد .

وفي حقل بيانات العنوان <titleStmt> تبدو المعلومات عن النص الإلكتروني والمصدر غير الإلكتروني متداخلة؛ فالعناصر الفرعية في هذا الحقل هي : العنوان <title> ، الذي يخصص لتسجيل عنوان العمل، واسم المؤلف <author> ، الذي يسجل به اسم مؤلف العمل (الأصلي) ، واسم الراعي <sponsor> ، والممول <funder> ، ومدير <principal> هؤلاء الرعاة والممولين ، والباحثون الرئيسيون المسؤولون عن إنشاء النسخة الإلكترونية من النص ، ثم يأتي بعد ذلك العنصر الفرعي الخاص ببيان المسؤولية <respStmt> . ومن الملاحظ أن عنصر بيان المسؤولية <respStmt> يتم في سياقه تحديد طبيعة المسؤولية وهوية المسؤول عن إعداد النسخة الإلكترونية للنص ، ومن ثم فقد يشتمل على أسماء الأشخاص ، والجهات المسؤولة عن تجميع العمل ، أو استنساخه ، أو تحويله إلى نسخة رقمية أو رقمنته ، أو ترميزه ، أو حتى المساهمة في إنشاء النص. وفي سياق عنصر بيان المسؤولية <respStmt> تتضمن العناصر الفرعية الخاصة بالأسماء <name> أسماء الأشخاص المسؤولين مسؤولية أساسية ، في حين تشتمل العناصر الفرعية الخاصة بطبيعة المسؤولية <resp> على توضيح علاقة هؤلاء الأشخاص الرئيسيين بالنص. على أن يتم ذلك على النحو التالي:

```
<titleStmt>
        <title>Wind and water : poem</title>
```

```
        <author>Walters, Winston</author>
        <respStmt>
                <resp>Creation of the machine-readable text:</resp>
                <name>Digital Library center, Midtwon
        University</name>
                <resp>TEI markup</resp>
                <name>John Monroem, Midtwon University</name>
        </respStmt>
</titleStmt>
```

وكما هو الحال بالنسبة للحقل الثاني من القواعد الأنجلو أمريكية للفهرسة، يشتمل بيان الطبعة <editionStmt> هنا في ترويسة مبادرة ترميز النص على المعلومات الخاصة بطبعة معينة من النص ، وعلى بيان المسؤولية الخاص بتلك الطبعة(١). لكن يجب ملاحظة أن بيان الطبعة <editionStmt> الذي يوجد في حقل العنوان <titleStmt> يشير إلى الطبعة الإلكترونية ، مع ملاحظة عدم استخدامه على نحو متكرر لعدم وجود إرشادات كافية يعتمد عليها كأساس للحالات التي تعامل فيها الإصدارات الإلكترونية المختلفة على أنها طبعات مختلفة.

ثم يأتي بعد ذلك حقل الحجم <extent> الذي يخصص لتسجيل الحجم التقريبي للملف الإلكتروني، هكذا: <extent>ca.٩٨٧٦ Kilobytes</extent>

.

(١) يعرف في اصطلاح الفهرسة ببيان الطبعة التبعي [المترجم].

أما حقل بيانات النشر <publicationStmt> فيصف تفاصيل نشر الملف الإلكتروني وتوزيعه، وقد يأتي هذا الوصف في شكل جمل نصية، أو قد يكون في عناصر فرعية، وفي هذا الحالة ينبغي أن يشتمل على عنصر واحد على الأقل من العناصر الفرعية الثلاثة التالية: الناشر <publisher> ، أو الموزع <distributor>، أو الجهة الرسمية <authority> . ويقصد بالجهة الرسمية الشخص أو الجهة المسؤولة عن إتاحة الملف الإلكتروني بأساليب أخرى مختلفة عن أساليب النشر أو التوزيع المعروفة . وقد يشتمل أي من تلك العناصر الفرعية الثلاثة على عناصر البيانات التالية : مكان النشر <pubPlace> لتحديد المكان الذي نشر أو وزع فيه العمل، وعنوان الناشر أو الموزع <address> ، ورقم التعريف <idno> الذي يستخدم لتحديد هوية العمل ، وكيفية الإتاحة <availability> لضوابط إتاحة العمل مثل قيود الاستخدام ، ثم تاريخ الإتاحة <date>. ويتم ذلك هكذا على النحو التالي :

```
<publicationStmt>
        <publisher>Midtwon University Electronic Texts
Project</publisher>
        <address>
                <addrLine>Box ٥٧٠</addrLine>
                <addrLine>Midtwon University</addrLine>
                <addrLine>Midtwon, OH</addrLine>
        </addrLine>
        <date>١٩٩٩</date>
        <idno type="MUETP">٩٩-١٨٢</idno>
</publicationStmt>
```

أما عنصر بيانات السلسلة <seriesStmt> فيستخدم في حالة صدور

المطبوع الإلكتروني ضمن حلقات سلسلة ما ، إذ تسجل فيه بيانات تلك السلسلة التي قد تحتوي على العناصر الخاصة بعنوان السلسلة <title>، ورقم هذا العمل فيها <idno> ، وبيانات المسؤولية الخاص بالسلسلة <respStmt>. ومن الملاحظ أن عنصر رقم التعريف <idno> يستخدم كذلك للأرقام التي تعرّف بالسلسلة مثل: الرقم الدولي المعياري للدورياتISSN [١].

أما حقل التبصرات <notesStmt> فيخصص لأية معلومات أخرى لها علاقة بالنص الإلكتروني بيد أنها لم ترد ضمن أي من الحقول المعرفة السابقة ، على أن ترد كل معلومة في تبصرة كعنصر فرعي مستقل، وذلك ضمن تاج مستقل هو تاج التبصرة <note> .

وأخيراً يأتي العنصر الخاص بوصف المصدر <sourceDesc> ، وهو عنصر إجباري يصف المصدر أو المصادر التي اشتق منها الملف الإلكتروني. وقد يشتمل على وصف إنشائي بسيط، أو على إرجاعة ببليوجرافية مقننة ترد في العنصر الفرعي المسمى <bibl>، أو على وصف ببليوجرافي مقنن كامل ، وفي هذه الحالة يسجل في العنصر الفرعي المسمى <biblfull> الخاص بالبيانات الببليوجرافية الكاملة، وتجدر الإشارة أخيراً إلى أن جميع الحقول الفرعية المسموح بتضمينها في القسم الفرعي الخاص بوصف الملف <fileDesc>، يمكن تضمينها في عنصر البيانات الببليوجرافية الكامل <biblFull>. مثال ذلك ما يلي:

<source Desc>

(١) هذا الوضع مخالف لما تنص عليه القواعد الأنجلو أمريكية للفهرسة ، حيث يسجل هذا الرقم في حقل مستقل وليس في حقل السلسلة [المترجم].

```
        <biblfull>
                <titleStmt>
                        <title>the Unmaking of a Mayor</title>
                        <author>William F. Buckley,Jr.</author>
                </titleStmt>
                <publicationStmt>
                        <publisher>the Viking Press, Inc.</publisher>
                        <pubplace>New York</pubplace>
                        <date>١٩٦٦</date>
                        <idno type="LCCN">٦٦-٢٠٣٣٩</idno>
                </publicationStmt>
        </biblfull>
</source Desc>
```

وصـف الخـصائص غيـر الببليوجـرافية أو وصـف البـروفايل profile Description :

إذا كـان قسم وصـف الملـف <fileDesc> يأتـي مناظـراً لمـا يتم فـي سياق الفهرسة الوصفية، فإن قسم وصف البروفايل <profileDesc> سيكون مناظراً تقريباً للفهرسة الموضـوعية ، علـى الأقل فـي محتوى العنـصر الاختـياري الخـاص بتـصنيف الـنص <textClass> ، والـذي يشتمل على عنصر واحد أو أكثر من العناصر التالية:

- عنـصر <keywords> الـذي يخصص للكلمـات الأساسية أو العبارات التـي تعرف بموضـوع النص أو طبيعته.
- عنـصر <classCode> الـذي يخـصص لرمـز التـصنيف الـذي يعطـى

للنص.

- عنصر <catRef> الذي يخصص للفئات الفرعية داخل بعض نظم التصنيف.

مع ملاحظة أن التاجين الخاصين بكل من الكلمات الأساسية <keywords> ، ورقم التصنيف <classCode> ، يأخذان الخاصية المقننة المسماة " خطة scheme " لتحديد كل من اللغة المقيدة وخطة التصنيف المستخدمتين في هذا الصدد . هكذا :

```
<textClass>
        <keywords scheme="LCSH">
                <term>Written communication-Social aspects</term>
                <term>Communication and technology</term>
                <author>William F. Buckley,Jr.</author>
        </keywords>
</textClass>
```

وإذا حذفت تلك الخاصية - أي scheme - فيفترض استخدام الكلمات الحرة أو غير المقيدة ، أما إذا ذكرت فإن قيمتها يجب أن تحدد مسبقاً في عنصر بيانات التصنيف <classDecl> بقسم وصف الترميز <encodingDesc> ، كما سيتضح في فقرات لاحقة.

وإضافة إلى عنصر تصنيف النص <textClass> ، فإن قسم وصف البروفايل <profileDesc> يمكن أن يشتمل على عنصري إنشاء العمل <creation> للمعلومات الخاصة بإنشاء النص ، واللغة المستخدمة

<langUsage> للمعلومات عن اللغات واللهجات المستخدمة في النص.

وصف الترميز **Encoding Description** :

مع أن ترويسة مبادرة ترميز النص قصد من ورائها أن تقدم فئات الوصف الببليوجرافي والموضوعي نفسها التي يستخدمها المفهرسون في المكتبات ، إلا أنه يضاف إلى ذلك أنها تسعى إلى القيام ببعض الوظائف غير الببليوجرافية ؛ من هنا يهدف قسم وصف الترميز <encodingDesc> إلى بيان الأساليب المتبعة في إنشاء النص الإلكتروني وترميزه بالتفصيل ، وكأنه بذلك أشبه بما تقوم به " أدلة الترميز [١] والأدلة التقديمية المصاحبة عادة لمجموعات البيانات الإلكترونية" .

ويشتمل قسم وصف الترميز <encodingDesc > على العناصر التالية:

- عنصر وصف المشروع <projectDesc> ، ويصف مشروع إنشاء الملف الإلكتروني ، بما في ذلك الهدف من إنشاء الملف ، وكيفية إنشائه.
- عنصر وصف العينة أو عملية المعاينة <samplingDesc> ، ويصف المبررات الأساسية والأساليب التي اعتمد عليها في اختيار عينات النص التي يتم تضمينها في المجموعة عند إنشائه.

---

(١) أدلة الترميز أو التكويد عبارة عن وثيقة تشتمل على معلومات عن بنية ملف بيانات العلوم الاجتماعية ومحتواه وإخراجه، أو دليل يصنف تنظيم ومحتوى ملف بيانات أو أكثر ، كذلك قائمة بملفات أو مصطلحات مع مقابلاتها في شكل مكود [المترجم].

- بيانات التحرير <editorialDecl> ، ويصف الأساسيات المتبعة في تحرير النص.

- بيانات التيجان <tagsDecl> ، ويعطي قائمة بالتيجان المستخدمة في ترميز النص ، مع إرشادات استخدامها.

- بيانات المراجع <refsDecl> ، ويحدد الكيفية التي أعدت بها قائمة المراجع.

- بيانات التصنيف <classDecl> ، وتسجل فيه نظم التصنيف taxonomies التي اعتمد عليها في العمل ، أو في اختيار المصطلحات الموضوعية ، أو في كليهما.

ويعد عنصرا بيانات التحرير <editorialDecl> ، وبيانات التيجان <tagsDecl>. من أكثر العناصر أهمية للباحثين الذين يستخدمون النص ؛ حيث يستخدم أولهما - أي بيانات التحرير <editorialDecl> - للإشارة إلى أية تغييرات وقرارات اتخذت تجاه سياسة تحرير النص ، كالإشارة مثلاً إلى إلغاء الشَّرطة في نهاية السطور، واستخدام علامات الهجاء والترقيم في شكلها المعتاد ، وأنماط الترميز المستخدمة ، وإذا ما كانت عملية الترميز أجريت يدوياً أم آلياً ، ومدى تدقيق النص لغوياً ، وما الطريقة المتبعة ، وغير ذلك من معلومات ذات علاقة بهذا الجانب.

أما عنصر بيان التيجان <tagsDecl> فيقدم معلومات تفصيلية عن عملية الترميز المعتمدة ، والكيفية التي ينبغي أن يعرض بها الترميز أو أن يترجم بها.

وإذا ما استخدم هذا العنصر فينبغي أن يضمن بداخله كل تاج معتمد في عملية الترميز، ويعطى داخل هذا العنصر (أي <tagsDecl>) العنصر الفرعي الخاص باستخدام التاج <tagUsage> الذي يشير إلى اسم التاج محدداً بالخاصية المسماة "gi" ، وعدد مرات تكرار العنصر في النص محدداً بالخاصية "occurs". هكذا:

```
<tagsDecl>
        <tagUsage gi="p" occurs="١٠١">
<tagsDel/>
```

أما العنصر الفرعي الخاص بتفسير النص أو ترجمته <rendition> فيشير إلى الكيفية التي ينبغي أن يعامل بها محتوى العمل ، وهذا العنصر الفرعي له علاقة بالعناصر التي تحددها الخاصية المسماة "render ". [هكذا :[

```
<tagsDecl>
    <rendition id ="rend.p">break indent</rendition>
    <tagUsage gi ="p"occur="١٠١"rende="rend.p"</rendition>
</tagsDecl>
```

ويعرّف عنصر بيانات التصنيف <classDecl> بخطط التصنيف أو اللغات المقيدة المستخدمة في قسم وصف البروفايل من الترويسة. ويمكن تعريف نظام التصنيف عن طريق الإحالة إلى بعض الخطط المعرفة خارجياً ، أو يمكن وصفه مباشرة داخل عنصر بيانات التصنيف باستخدام العنصرين الفرعيين: الفئة <category> ، ووصف الفئة <catDesc>. وفي

كلتا الحالتين يشتمل العنصر الفرعي الخاص بالبيانات الببليوجرافية المسمى <bibl> على إرجاعة ببليوجرافية غير مقننة لاسم الخطة. [هكذا]:

```
<classDecl>
        <taxonomy id="DDC">
            <bibl> Dewy Decimal System</bibl>
        </taxonomy>
</classDecl>
        <taxonomy id="MT">
                <bibl>Midtwon Local Classification Scheme</bibl>
                <category id="MT.١٠٠">
                        <catDesc>University Administration</catDesc>
                </category>
                <category id="MT.١٠٥">
                        <catDesc> Office of the President</catDesc>
                </category>
                <category id="MT.١١٠">
                        <catDesc> Office of the Provest</catDesc>
                </category>
                …..
                </taxonomy>
</classDecl>
```

وصف المراجعة أو التنقيح **Revision Description** :

وأخيراً يشتمل قسم وصف المراجعة <revisionDesc> على تاريخ المراجعات التي أجريت على النص الإلكتروني ، ويمكن تسجيل

الملاحظات الخاصة بهذا الأمر في شكل نص غير بنائي داخل عنصر التغييرات <change> ، أو في شكل نص بنائي باستخدام العناصر الفرعية التالية : تواريخ المراجعة ، وأسماء المراجعين ، وطبيعة المراجعة ، والتغييرات التي أجريت. [هكذا]:

```
<revisionDesc>
        <change>
                <date>August ١٩٩٨</date>
                <respStml>
                        <name>Emily Hart</name>
                        <resp>ed.</resp>
                </respStml>
                <item>Corrected tagging of proper names</item>
                <item>Completed taxonomy description in
        header</item>
        </change>
</revisionDesc>
```

ترويسة مبادرة ترميز النص وفهرسة المكتبات :

بالرغم من أن واضعي مسودة مواصفة ترويسة مبادرة ترميز النص توقعوا أن يتم إنشاء معظم الترويسات من قبل الباحثين في مجال الإنسانيات؛ إلا أن الواقع العملي أثبت أن معظم النصوص المرمزة وفقاً لمبادرة ترميز النص قد أنشئت في المكتبات أو في مراكز النصوص الإلكترونية المرتبطة بالمكتبات. وقد أشار إلى ذلك براد إيدن Brad Eden بقوله:

" ليس غريباً على الإطلاق أن تُنشأ الترويسة من أجل المفهرس، ولكن الغريب أن ينشئها المفهرس، ومن الغريب كذلك أن تحول مسودة الترويسة

إلى صيغة مارك ، وتضاف مباشرة إلى النظام المحلي للفهرسة حيث تُجرى عليها التعديلات النهائية ، ثم تحول مرة أخرى من صيغة مارك إلى صيغة لغة الترميز المعيارية العامة ، ثم يعاد تضمينها في الوثيقة(١).

وفي بعض المشروعات تستخدم تسجيلة مارك للنسخة المطبوعة من أحد الأعمال كأساس لإعداد ترويسة مبادرة ترميز النص، إلا أنه ـ وبصرف النظر عن الشكل الذي يتدفق عبره هذا النشاط ـ من الشائع في سياق كل من الفهرسة التقليدية للمكتبات ومبادرة ترميز النص أن يتم هذا الإجراء بالنسبة للعمل نفسه، لأن ترويسة مبادرة ترميز النص تكون مرمزة في صيغة لغة الترميز المعيارية العامة كما هو الحال بالنسبة للنص الذي تصفه، فيمكن بحثها وعرضها عن طريق نظم النصوص الكاملة نفسها المستخدمة في تداول نصوص مبادرة ترميز النص TEI texts. غير أن المكتبات تفضل أن يكون لديها التسجيلات الببليوجرافية لجميع مقتنياتها المعروضة في فهارسها المباشرة، مما يتطلب معه إنشاء تسجيلة مارك مستقلة لأي من هذه المقتنيات ليتم إدخالها في منظومة الفهرس، ونتيجة لذلك يعنى كل من مجتمعي الفهرسة ومبادرة ترميز النص بدراسة طبيعة العلاقة بين هذين النمطين من خطط ما وراء البيانات(٢).

(١) Brad Eden, "Metadata, TEl, and the Academic Library Community: An Update," in <Tel ١٠>: Text Encoding Initiative Tenth Anniversary User Conference, available at http://www.stg.brown.edu/conferencesltei١٠/tei١٠/tei١٠.papers/eden.html. Accessed ١٩ June ٢٠٠٢.

(١) يقصد تسجيلة مارك ، وتسجيلة ترويسة مبادرة ترميز النص [المترجم].

ومن جانب الفهرسة قامت لجنة المهام التابعة للجنة الفهرسة CC:DA(١) بتقييم ترويسة مبادرة ترميز النص كمصدر لما وراء بيانات تسجيلات الفهرسة ، وكانت لجنة المهام الخاصة بما وراء البيانات وقواعد الفهرسة قد كونت لدراسة مدى استخدام خطط ما وراء البيانات غير التقليدية أو التسجيلات المشتقة عنها في فهارس المكتبات. وخلص تقريرها النهائي المرتكز أساساً على دراسة كل من ترويسة مبادرة ترميز النص ودبلن كور، إلى أن ما وراء البيانات لا يمكن دمجها في فهارس المكتبات ما لم تكن قد أنشئت وفقاً للقواعد الأنجلو أمريكية للفهرسة في وصفها المادي ، وعلى إحدى اللغات الموضوعية المقيدة كقائمة رؤوس موضوعات مكتبة الكونجرس في وصفها الموضوعي(٢).

وهناك تقرير فرعي أكثر تفصيلاً عن ترويسة مبادرة ترميز النص أفاد بأن الترويسة قد حققت وظائف أكثر من الوظائف التي حققها الوصف الببليوجرافي ؛ وبالتالي فلا تقوم ترويسة مبادرة ترميز النص مقام تسجيلة الفهرسة التقليدية أو العكس ، كما لا تعد أياً منهما بديلاً عن الأخرى ، ثم أشار التقرير إلى عدم تناغم أو توافق محتوى مبادرة ترميز النص مع مصطلحات القواعد الأنجلو أمريكية للفهرسة، ومن هذه التناقضات أن منشئ الترويسة في الأغلب لا يكون ملماً بالقواعد الأنجلو أمريكية للفهرسة ، كما أن المصادر التي تستقى منها معلومات

---

(١) هي لجنة تابعة للجمعية الأمريكية للمكتبات مهمتها متابعة قواعد الفهرسة بالتعديلات ، وتسمى لجنة الفهرسة : Cataloging Committee : Description and Access - انظر الفصل السادس [المترجم].

(٢) Committee on Cataloging: Description and Access, Task Force on Metadata and the Cataloging Rules, Final Report (August ٢١،١٩٩٨), available at http://www.ala.org/aicts/organization/ccs/ccda/tf-tei٢.html. Accessed ١٩ June ٢٠٠٢.

الوصف التي قررتها القواعد الأنجلو أمريكية للفهرسة لم تتبع في مبادرة ترميز النص، كذلك لم تتبع قواعد كتابة الحروف الاستهلالية وعلامات الترقيم، كما لم تستخدم الأشكال الاستنادية الخاصة بالأسماء التي أقرتها تلك القواعد . وباختصار فإن مجمل خلاصة التقرير والتي لم تأت مفاجئة تفيد بأنه إذا لم يتبع منشئو ما وراء البيانات قواعد الفهرسة الخاصة بالمحتوى والتي تقررها القواعد الأنجلو أمريكية للفهرسة ، فمن غير المحتمل أن تأتي ما وراء البيانات متوافقة مع تلك القواعد.

وفي الوقت نفسه الذي نشطت فيه لجنة المهام - CC:DA - عقد مؤتمر في مكتبة الكونجرس حول " مبادرة ترميز النص ولغة الترميز الموسعة في المكتبات الرقمية TEI and XML in digital libraries " ، وركزت إحدى المجموعات التي أدارت دفة المناقشات في اللقاء على "ما وراء البيانات الوصفية: مارك والقواعد الأنجلو أمريكية للفهرسة وترويسة مبادرة ترميز النص "، واتفق أعضاء المجموعة على أن هناك حاجة لإيجاد " قابلية تحويل convertibility " جيدة بين كل من الترويسة ومارك ، كما قدموا توصية تعنى بالمحتوى جاءت على النحو التالي :

- السعي للتوصل إلى اتفاق حول أفضل الممارسات ، وإعداد دليل إرشادي شامل toolkit لتطبيق ترويسة مبادرة ترميز النص بحيث يشمل ما يلي:

- إرشادات البيانات الوصفية - القواعد الأنجلو أمريكية للفهرسة المبسطة (?) AACRlite ، والقوائم الاستنادية.

- إرشادات عمل التيجان .

- مواصفات الحد الأدنى من عناصر البيانات الوصفية.

- إرشادات اختيار النص(١).

ومتابعة لهذه الجهود، اجتمع ممثلون عن المبادرات الرئيسية لترميز النصوص من جامعتي ميتشيجن وفرجينيا لمراجعة طبيعة العلاقة بين ترويسة مبادرة ترميز النص ومارك ، واقتراح أفضل الممارسات لإنشاء ترويسات مبادرة ترميز النصوص للوصول إلى "المحتوى المتوافق compatible content ". كما انتهوا في تقريرهم إلى تحديد مجموعة مصادر المعلومات الرئيسية التي ينبغي أن تستخدم مع الأشكال المختلفة للنصوص الإلكترونية ، بالإضافة إلى تقديم إرشادات اختيار محتوى جميع التيجان ، حيث تضمنت الإرشادات الخاصة باستخدام عنصر المؤلف <author> في قسم وصف الملف <fileDesc> التوجيه التالي :

" استخدم تيجانًا مستقلة مع تاج المؤلف <author> لكل من الاسم الأخير، والاسم الأول، والاسم الأوسط ، والتاريخ ، واللقب الوظيفي، لإتاحة مزيد من المرونة عند عرضها وتكشيفها وتحويلها إلى مارك. كذلك قم بإعداد أشكال ثابتة على المستوى الوطني للأسماء كلما أمكن ذلك، على أن

---

(١) TEl and XML in Digital Libraries, June ٣٠-JuIy I, ١٩٩٨, Washington, D.C., Working Group ١: Descriptive Metadata: MARC, AACR٢, and the TEl Header Discussion Summary and Recommendations, available at http://www.umdl.edulworkshops/teidIf/teigrpl.html. Accessed ١٩ June ٢٠٠٢.

يرد الاسم في صيغته المقلوبة ، وأن يتم إدخاله بالشكل الثابت"([1]).

ومن الملاحظ أن معظم تلك الإرشادات لها أهميتها حتى خارج مجال ربطها بفهرسة المكتبات ، كما أن ترويسة مبادرة ترميز النص نفسها مثلها تماماً مثل إرشادات مبادرة ترميز النص قد أعدتا سوياً بشكل يتيح للقائم بالترميز أكبر قدر من المرونة فيما يتصل بنوعية البيانات التي يجب تسجيلها ، وبكيفية تتويج هذه البيانات وفق صيغة لغة الترميز المعيارية العامة . ومع أن هذه المرونة تدفع لتوسعة اتجاهات تطبيق هذه الترويسة ، إلا أن ما يصاحب ذلك من تفاوت يمكن أن يعيق إمكانات تبادل أو تقاسم أو بحث تلك الترويسات التي أعدت من قبل مشروعات مختلفة. ومع ذلك يمكن القول إن الإرشادات يمكن أن تتكفل بالتغلب على مثل هذا التفاوت عن طريق تشجيع الحرص على التناغم بين الممارسات عند إنشاء كل من المحتوى والترميز؛ حيث توضح هذه الإرشادات - على سبيل المثال - العناصر التي ينبغي استخدامها لوصف المصدر الأصلي ، كذلك المصادر التي ينبغي استخدامها لوصف النص الإلكتروني، كما تقترح أشكالاً مقننة لبعض العناصر كعنصري الحجم <extent> ، والتاريخ <date> ، بالإضافة إلى توضيحها لطبيعة العلاقة بين التيجان المختلفة.

لقد باتت ترويسة مبادرة ترميز النص من خطط ما وراء البيانات واسعة

(٢) TEl/MARC "Best Practices" (June ١٦،٢٠٠١), available at http://www.lib.umich.edu/staff/ocu/teiguide.html. Accessed ١٩ June ٢٠٠٢.

الاستخدام ، وتم تكييفها لتناسب مجموعة كبيرة من المواد النصية المرمزة وفقاً لأي من لغة الترميز المعيارية العامة ولغة الترميز الموسعة التي تخرج عن إطار النصوص العلمية في الإنسانيات، بما في ذلك مقالات الدوريات ، ومقالات الصحف، والرسائل الأكاديمية الإلكترونية ، ونظراً لإمكانية تضمين الترويسة في الوثيقة التي تصفها ، فقد أصبحت مناسبة للتطبيقات التي تقوم بالبحث في كل من البيانات وما وراء البيانات على حد سواء ، كما أنها أصبحت مناسبة كذلك للتطبيقات التي تحتاج إلى كيانات ذاتية التوثيق self-documenting objects.

وبغض النظر عما إذا كانت الأجزاء الببليوجرافية من الترويسة قد جاءت متوافقة مع القواعد الأنجلو أمريكية للفهرسة ومارك، فإن أهمية ترويسة مبادرة ترميز النص في توثيق الجوانب غير الببليوجرافية للنص، بما في ذلك تفاصيل ترميزه ومراجعته، تكفل لها أنها ستظل هي المعيار السائد لوصف النصوص الإلكترونية، يضاف إلى ذلك أن هذه الترويسة اعتبرت نموذجًا لمكون الوصف الببليوجرافي لخطط ما وراء بيانات أخرى متعددة قائمة على لغة الترميز المعيارية العامة ولغة الترميز الموسعة ، بما في ذلك الوصف الأرشيفي المرمز "encoded Archival Description "EAD ، ومبادرة توثيق البيانات"Data Documentation Initiative "DDI .

READINGS :

- Committee on Cataloging: Description and Access, Task Force on Metadata and the Cataloging Rules. Final Report (August ٢١,

١٩٩٨). Available at
http://www.ala.org/alcts/organization/ccslccda/tf-tei٢.html.
Includes a chapter analyzing the TEl header for its usability in library systems.

- Seaman, David. The Electronic Text Center Introduction to TEl and Guide to Document Preparation. Available at http://etext.lib.virginia.edultei/uvatei.html. The University of Virginia's internal guidelines for marking up documents according to TEl. Includes a chapter on the TEl header and examples of completed headers for various types of materials.
- TEl and XML in Digital Libraries. Available at http://www.umdI.umich.edulworkshops/teidIfI.
  Notes from a meeting at the Library of Congress, held June ٣٠-July ١, ١٩٩٨; includes consideration of compatibility between MARCI AACR٢ and TEl.
- Text Encoding Initiative (home page). Available at http://www.tei-c.org/. Accessed ٢٠ June ٢٠٠٢.
  Information about the TEl consortium, history, projects, and Guidelines.

الفصل الثامن

دبلــن كــور

مجموعة دبلن كور لعناصر ما وراء البيانات خطة عامة الأغراض لوصف مصادر المعلومات وتهدف في الأساس إلى تيسير استكشاف كيانات المعلومات المتاحة على الوب.

مجموعة العناصر :

لا يزال الطابع الأسطوري يحيط بأصل دبلن كور حتى الآن ، مع أن البداية تعود إلى خريف عام ١٩٩٤م حيث عقد المؤتمر الدولي الثاني للشبكة العنكبوتية العالمية في شيكاغو حول " متصفح موزايك والوب Mosaic and the Web". ومع أن معظم محاور ذلك المؤتمر اتجهت لإقرار إمكانية الوب الناشئة [آنذاك] على تغيير الطريقة التي تقدم بها المعرفة، فقد تركز جل اهتمام المشاركين على دراسة الكيفية التي يمكن عن طريقها التعرف إلى هذا المحتوى المتاح حديثاً على تلك الشبكة، وقد بادر ثلاثة من الحضور - هم "ستو ويبيل Stu Weibel" من مركز الحاسب الآلي للمكتبات على الخط المباشر (أو سي إل سي OCLC) ، وجوزيف هاردين Joseph Hardin من NCSA ، والراحل يوري روبينسكي Yuri Rubinski من سوفتكواد Softquad - إلى الدعوة إلى عقد ورشة عمل متعددة التخصصات في مارس عام ١٩٩٥م لتحديد عناصر البيانات الوصفية الضرورية لاستكشاف مصادر المعلومات الشبكية، ورؤي أن يخصص اللقاء من أجل " التوصل إلى اتفاق على مجموعة رئيسية من عناصر البيانات للنظائر الوثائقية، واستكشاف عناصر البيانات لغير ذلك من الأنواع الأخرى من الكيانات الشبكية" .

ونظراً لانعقاد ورشة العمل هذه في المقر الرئيسي لمركز (أو سي إل سي) في دبلن بولاية أوهايو، ارتبط اسم هذا المدينة - أي دبلن - بعنوان المجموعة الرئيسية من عناصر البيانات المقترحة - أي دبلن كور أو مجموعة دبلن الأساسية Dublin Core "، كما أن ورشة العمل نفسها سميت "دبلن كور الأولي DC١ " إشارة إلى الحلقة الأولى من سلسلة ورش العمل المستمرة التي عقدت في هذا الخصوص.

أما جهود تطوير المواصفات الرسمية لدبلن كور فتديرها هيئة تعرف "بمبادرة دبلن كور لما وراء بيانات Dublin Core Metadata Initiative "DCMI"" ، تتألف من إدارة صغيرة تعمل بأجر لها مجلس أمناء استشاري مع مجموعة كبيرة من المتطوعين. وبمرور الوقت طورت هذه المبادرة بنية حكومية وأعدت مجموعة من الإجراءات الرسمية لإقرار المواصفات الجديدة وما يستجد من مصطلحات ، ورأت أن تنفذ أنشطتها من خلال مجموعات عمل تلتقي عبر منتديات لمناقشة قضايا محددة وإعداد مسودات وثائق المتطلبات والمواصفات، على أن تنتهج في سبيلها لإقرار المواصفات نهجاً مشابهاً لما يقوم به ائتلاف الوب "W٣C" الذي يعتمد على أسلوب تطور الحالة progression of statuses ، حيث تبدأ بموجب هذا الأسلوب رحلة إقرار المواصفة بمرحلة تقديم "المسودة" ، ثم تقديم " التوصيات أو المقترحات عليها " ، ثم إقرار التوصية النهائية التي تقابل المواصفة القياسية ، وعندها تعتبر المواصفة مستقرة ، وللمنفذين أن يتبنوا تطبيقها . أما بالنسبة لإقرار مصطلحات الجديدة لما وراء البيانات (سواء كانت خاصة بعناصر

البيانات elements أو بالمقيدات qualifiers) فتسند مسؤوليتها للجنة صغيرة عالية المستوى تعرف " بلجنة الاستخدام usage board ".

تتكون مجموعة دبلن الأساسية لعناصر ما وراء البيانات (أو اختصاراً دبلن كور) من خمسة عشر عنصراً من عناصر البيانات. نوردها في الفقرات التالية - كما اقتبسناها من أحد مصادرها الرسمية(١) - في شكل يبين بالنسبة لكل عنصر منها اسمه (أي المعرف Identifier) ، وتعريفه (أي تحديد مجال استخدامه definition):

- المعرف : العنوان .

التعريف : اسم المصدر .

- المعرف : المنشئ .

التعريف : الكيان المسؤول مسؤولية أساسية عن محتوى المصدر .

- المعرف : الموضوع .

التعريف : الموضوع الذي يدور حوله محتوى المصدر .

- المعرف : الوصف .

التعريف : وصف محتوى المصدر .

- المعرف : الناشر .

(١) Dublin Core metadata Element Set, Version ١,١: Reference Description, available at http://dublincore.org/documents/dces/. Accessed ٢١ June ٢٠٠٢.

التعريف : الكيان المسؤول عن إتاحة المصدر .

- المعرف : المساهم أو المشارك .

التعريف : الكيان المسؤول عن تقديم مساهمات تتصل بمحتوى المصدر .

- المعرف : التاريخ .

التعريف : تاريخ مرتبط بحدث ما في دورة حياة المصدر .

- المعرف : النوع .

التعريف : طبيعة محتوى المصدر .

- المعرف : الشكل .

التعريف : الشكل المادي أو الرقمي الذي يعرض به المصدر .

- المعرف : المعرف .

التعريف : إشارة صريحة للمصدر عندما يرد في سياق معين .

- المعرف : المصدر .

التعريف : إشارة إلى المصدر الذي أخذ منه المصدر الحالي .

- المعرف : اللغة .

التعريف : لغة المحتوى الفكري للمصدر .

- المعرف : العلاقة .

التعريف : إشارة إلى مصدر آخر ذي علاقة بالمصدر الحالي .

- المعرف : التغطية .

التعريف : المجال الذي يغطيه محتوى المصدر .

- المعرف : الحقوق .

التعريف : معلومات عن الحقوق المقررة للمصدر أو المترتبة على استخدامه.

ومما تجدر الإشارة إليه في هذا الصدد أن جميع هذه العناصر اختيارية، كما أنها قابلة للتكرار، يضاف إلى ذلك أن الخطة نفسها تتسم باستقلالية الشكل أو الصيغة format-independent ، بمعنى أنها لا ترتبط بأي من طرق تمثيل البيانات ، كما هو الحال - مثلاً - مع ترويسة مبادرة ترميز النص التي ترتبط بكل من لغة الترميز المعيارية العامة ولغة الترميز الموسعة ، كما أنها غير مرتبطة كذلك بأي مجموعة معينة من قواعد المحتوى ، وإن كانت هناك إشارة ترد في خاصية التعليقات الخاصة ببعض العناصر إلى أفضل الممارسات المقترحة، ناهيك عن بعض التوصيات الأخرى الإضافية التي ترد في سياق أحد أدلة الاستخدام الرسمية(١). وباتباع تلك التوصيات ، يبدو وصف دبلن كور المبسط هكذا على النحو التالي:

---

(١) Diane Hillmann, Using Dublin Core, available at http://dublincore.org/٢٠٠١/٠٤/١٢/usageguide/ Accessed ٢١ June ٢٠٠٢.

Title= " the electronic text center introduction to TEI and Guide to Document Preparation"

Creator="Seaman, David "

Subject=" Text encoding Initiative"

Subject=" SGML markup rules"

Description="Guidelines written by University of Virginia Electronic Text Center for markup electronic texts using the TEILTE.DTD, a subset of the TEI tagset "

Date=" ١٩٩٥"

Type="text "

Language="en "

Identifier=" http://etext.lib.virginia.edu./tei/uvatei.html"

وهنالك مقيدات دبلن كور Dublin Core qualifiers ، وهي عبارة عن مواصفة مصاحبة لمجموعة دبلن كور لعناصر ما وراء البيانات(١). وللمقيد وظيفتان: منها ما يقوم بتعريف خطة الترميز المعتمدة في تمثيل أي من عناصر دبلن كور، ومنها ما يركز على توضيح معاني عناصر البيانات. وإذا كان مقيد خطة الترميز encoding scheme qualifier هو الذي يشير إلى الخطة أو القائمة الاستنادية المعتمدة في تمثيل قيمة عنصر البيانات ، فإن مقيد توضيح العنصر element refinement qualifier يستطيع أن يحدد مجال معنى عنصر البيانات أو يضيِّقه، لكنه في الوقت نفسه لا يستطيع أن يوسع

(١) Dublin Core qualifiers, available at http://dublincore.org/documents/dcmes-qualifiers/. Accessed ٢١ June ٢٠٠٢.

هذا المعنـى أو يغيره. ولعل من أهم ما تتسم به هذه المقيدات الأخيرة - أي مقيدات توضيح العنصر - أنه يمكن تجاهلها ، ومع ذلك سيظل المعنى الذي تأخذه قيمة عنصر البيانات مفهوماً . ويعتمد هذا المطلب الذي يعرف بمبدأ " التجاهل dumb down principal"(١) على حقيقة مفادها أن ليس بالضرورة أن تتعـرف جمـيع تطبـيقات معالجـة مـا وراء بـيانات دبلـن كـور إلـى جمـيع المقيدات، ومن ثم يكون من الممكن لها " تجاهل [المقيد] والاعتماد علـى المعنى الأساسي غير المقيد لعنصر البيانـات.

وتجـدر الإشـارة كـذلك إلـى أن المقـيدات مخصـصة لعناصـر البـيانات الفـردية، فلعنـصر الـتاريخ - علـى سـبيل المـثال - خمـسة مقـيدات توضـيح معـتمدة (هــي : محــدث Created ، وصــالح Valed ، ومــتاح Available ، ومنشور Issued ، ومعدل Modified) . ومقيدان معتمدان من مقيدات خطة الترميز (مقيد مبادرة دبلن كور لخطـة ترميز الفترات الزمنية(٢) DCMI period ، وصـيغة ائتلاف الوب لترميز التاريخ والوقت W٣C-DTF)(٣) ، كذلك لعنصر البيانات مقيد توضيح واحد (هو : عنوان بديل) ، وليس له مقيد خطـة الترميز ، مع ملاحظـة أن للمقيدات تعريفاتها الخاصـة التي قد تحيل إلى مواصفات أو قوائم استناد أخرى.

---

(٢) في دبلن كور ، المبدأ الذي ينبغي بموجبه أن يكون عنصر البيانات المقيد مفهوماً للفرد أو للتطبيق اللذين لا يعرفان المقيد [المترجم].

(١) تشير بذلك إلى خطـة مبادرة دبلن كور لترميز الفترات الزمنية التي صدرت بشأنها مواصفة مستقلة بعنوان : DCMI Period Encoding Scheme: specification of the limits of a time interval, and methods for encoding this in a text string [المترجم].

(٢) يشير الاختصار إلى W٣C Date and Time Format [المترجم].

ومع ما تتسم به خطة دبلن كور من بساطة فقد ظهرت بعض المشكلات عند تطبيقها ، من هذه المشكلات تداخل المعاني عند تعريف بعض عناصر البيانات ، حيث يمكن أن يفهم عنصر " المنشئ " - مثلاً - على أنه المشارك ، كما قد يفهم عنصر " المصدر " على أنه أحد أشكال عنصر " العلاقة" ، وقد أحدث ذلك لبساً لدى مطبقي هذه الخطة ، وخاصة حول مدى اعتبار عنصر معين هو الأنسب للاستخدام من عنصر آخر . وإزاء هذه الإشكالات قدم اقتراح بأن تدمج هذه العناصر الثلاثة : "المنشئ ، والمشارك ، والناشر" في عنصر واحد يسمى " الوكيل agent " ، غير أن هذا الاقتراح لم يحظ بالقبول ورُفض ، وهو ما حدث بالنسبة للاقتراح الآخر الذي كان يقضي بالحد من استخدام عنصر "المصدر" لصالح استخدام عنصر "العلاقة" . ومع ذلك لا يزال كثير من مطبقي هذه الخطة ينظرون إلى صحة مبررات تضمين عنصري "المنشئ" و"المصدر" في المواصفة الأصلية ؛ فطالما دأب المجتمع الببليوجرافي على إعطاء حالة "التأليف " صفة الحالة الخاصة، كما ينعكس ذلك على مفهوم القواعد الأنجلو أمريكية للفهرسة للمدخل الرئيسي ، يبدو تمييز الدور الخاص للمنشئ عن أدوار المشاركين الآخرين منطقياً أو له ما يبرره عملياً في بعض التطبيقات. وقياساً على ذلك يتبين أن الاستخدام الهام لدبلن كور ينصب على وصف الإصدارات الإلكترونية الناشئة عن مشروعات التحويل الراجعة [للمصادر التقليدية]، وعلى ضوء ذلك يمكن تبرير وجود عنصر خاص لتسجيل معلومات عن المصدر غير الرقمي المناظر للمصدر الإلكتروني.

وهناك قضية أخرى هامة تتصل بطبيعة الوصف في حالة وجود عدة

إصدارات للمصدر الواحد ؛ حيث يسجل بشكل عام اسم المصور على أنه "منشئ" الصورة ، ولكن إذا حولت الصورة من شكلها العادي إلى الصيغة الرقمية، ولتكن صيغة جي بيج JPEG ، يصبح من الممكن القول بأن منشئ الصورة هو من قام بعملية المسح الضوئي ، لكن هناك من مطبقي دبلن كور من يرى أن الفني الذي يقوم بعملية المسح الضوئي لم يقم بجهد فكري، ومع ذلك يجب أن يسجل على أنه "مشارك" إذا ظهر اسمه في جميع الإصدارات ، بينما يرى آخرون أن تسجيلة دبلن كور يجب أن تصف المصدر المتاح بين أيدينا ، والذي يشير إلى أن المنشئ هنا بالنسبة للصورة الرقمية في صيغة (جي بيج) هو القائم بعملية المسح الضوئي scanner ، أما المصور في أحسن حالاته فما هو إلا "مشارك"، ويقرر هذا المبدأ - الذي يعرف بمبدأ " معالجة كل حالة بحالتها one- to- one" - بأنه إذا وجدت عدة إصدارات من المصدر الواحد ، ينبغي أن توصف كل إصدارة منها بشكل مستقل عن الإصدارات الأخريات.

التركيبات التي تصاغ فيها عناصر دبلن كور :

مع أن التوصيتين اللتين تعرفان بكل من دبلن كور وبمقيدات دبلن كور ، قصد منهما نقل دلالات عناصر دبلن كور دون غيرها ؛ فلابد أن يكون لخطة ما وراء البيانات على الأقل تمثيل تركيبي واحد أو أكثر syntactical representation شريطة أن يكون مقبولاً بشكل عام ، حتى تصبح هذه الخطة قابلة للتطبيق ميدانياً. ولعل من أولى مواصفات الترميز التي وصلت إلى درجة "التوصية" تلك المواصفة التي كانت موجهة للغة ترميز النص الفائق

HTML [١]؛ والتي تستخدم تاج الميتا <meta> مع خاصيتي "الاسم" ، و" المحتوى " بالصيغة النوعية التالية :

```
<meta name =" PREFIX.Element_name"
Content = "element_value">
```

ولأن عنصر البادئة في المثال السابق يبدو أساسياً حيث يستخدم للربط بمواصفة دبلن كور، فإنه يمثل في صيغة لغة ترميز النص الفائق عن طريق مجموعة من الخصائص لعنصر " الرابط " <link> هكذا:

```
<link rel="schema.PREFIX"
href=http://purl.org/dc/elements/١,١/
Title="Dublin Core Metadata Element Set, Version ١,١">
```

وتقرر مواصفة لغة ترميز النص الفائق استخدام الحرف الأول من كلمة Dublin وكلمة Core في شكلهما الكبير هكذا : "DC" كعنصر البادئة ، بحيث يمكن أن يعرض جزء من وصف وثيقة سيمان - التي عرضت في فقرة سابقة - في صيغة لغة ترميز النص الفائق على النحو التالي:

```
<link rel="schema.DC"
href="http://purl.org/dc/elements/١,١/"
title="Dublin Core Metadata Element Set, Version ١,١">
<meta name="DC.Title">
```

---

(١) John Kunze, Encoding Dublin Core in HTML (IETF RFC٢٧٣١), available at http//www.ietf.org/rfc/rfc٢٧٣.txt. Accessed ٢١ June ٢٠٠٢.

```
content=" the electronic text center introduction to TEI and Guide to
Document Preparation">
<meta name="DC.creator">
Content="Seaman, David ">
<meta name="DC.Identifier">
Content=" http://etext.lib.virginia.edu./tei/uvatei.html">
```

أما مقيد خطة الترميز فيتم تمثيله عن طريق تاج الميتا لخاصية " الخطة scheme" هكذا:

```
<meta name="DC.DC.Type"
Scheme="DCMITYPE"
Content="text">
```

كما يتم تمثيل مقيد توضيح العنصر في ترميز نقطي[1] dot" notation " ، هكذا على النحو التالي:

```
<meta name="DC.Date.created"
Content="١٩٩٥">
```

ومع أن تمثيل دلالات عناصر دبلن كور في صيغة لغة ترميز النص الفائق يتسم بالوضوح والسهولة ، ولا تعوقه عقبات خصوصاً في البيئات التي تسعى فيها محركات البحث المهيأة للإفادة من إمكانات تيجان الميتا <meta> في تنقيبها وتكشيفها لصفحات الوب، فإن ثمة بعض المحدوديات والمعوقات التي تحيط بذلك ، حيث تفضل التطبيقات التي تستلزم وجود

---

(١) يقصد بالترميز النقطي الفصل بين رمز العنصر واسمه ومقيده بالنقط، هكذا - على سبيل المثال - : Dc. Date. Created [المترجم].

تسجيلات ما وراء البيانات [خارجية] ، وليس ما وراء البيانات المضمنة في الوثائق ذاتها ، استخدام صيغة لغة الترميز الموسعة XML كتركيبة للتبادل ، يضاف إلى ذلك أن صيغة لغة ترميز النص الفائق لا تستطيع تمثيل المكونات معقدة التركيب بشكل كبير ،كما هو الحال عندما يتطلب الأمر تصنيف العناصر المكررة في مجموعات فرعية حتى تصبح أكثر وضوحاً من حيث معناها.

ويمكن كذلك تمثيل دبلن كور في صيغة لغة الترميز الموسعة ، وفي هذا الصدد طورت عدة خطط معتمدة على هذه اللغة لتطبيقات معينة لدبلن كور، من هذه الخطط إحدى الخطط المعتمدة للاستخدام مع تطبيقات مبادرة الأرشيفات المفتوحة لجمع ما وراء البيانات open archives initiative metadata harvesting applications ، وتزخر الصفحة الرئيسية لمبادرة دبلن كور DCMI بقائمة من الروابط التي تحيل إلى كثير من الخطط التي يدعمها مجتمع دبلن كور. وفضلاً عن ذلك قام مكتب المملكة المتحدة لشبكات المكتبات والمعلومات UKOLN بإصدار " الإرشادات العامة لتمثيل دبلن كور المقيد وغير المقيد في صيغة لغة الترميز الموسعة(١)، وتدعو هذه الإرشادات المنفذين إلى الإفادة من خدمة الاسم المميز بلغة الترميز الموسعة XML namespace ، لتعريف عناصر دبلن كور بشكل منفرد ، والتي يجب أن تمثل كعناصر لغة الترميز الموسعة XML elements.

---

(١) Andy Powell and Pete Johnson, "Guidelines for Implementing Dublin Core in XML. ١٢ March ٢٠٠٢, available at http://www.ukoln.ac.uk/metadata/dcmi/dc-xml-guidelines./ Accessed ٢١ June ٢٠٠٢.

وفي ضوء ذلك تبدو وثيقة سيمان - سابقة الذكر - بعد تمثيلها في صيغة دبلن كور المبسطة طبقاً لمواصفة مكتب المملكة المتحدة لمشابكة المكتبات والمعلومات على النحو التالي:

```
<?xml version="١,٠"?>
<metadata
Xmlns=http://myorg.org/myapp/
Xmlns:xsi=http://www.w٣.org/٢٠٠١/XMLSchema-instance
Xsi:schemaLocation="http:myorg.org/myapp.org/myapp/scema.xs"
Xmlns:dc="http://purl.org/dc/elements/١,١/">
<dc:title>the electronic text center introduction to TEI and guide to
document prepation</dc:title>
<dc:creator>Seaman,David</dc:creator>
<dc:identifier>http://etext.lib.verginia.edu/tei/uvatei.html</cd:identifier>
</metadata>
```

وتبدو خطة لغة الترميز الموسعة المستخدمة في هذا المثال خطة (خيالية) يشير إليها المحدد الموحد لموقع المصدر المسمى "http://myorg.org/myapp/schema.xsd" ، كما أن الاسم المميز نفسه بلغة الترميز الموسعة لمجموعة عناصر دبلن كور يشار إليه ببيان الاسم المميز الذي يستهل بالوسيمة "=xmlns:dc" ، في إشارة إلى أن أسماء عناصر دبلن كور سوف تكون مسبوقة بالحرفين "dc:" ، وأن تعريف عناصر دبلن كور سيوجد في الوثيقة الموجودة في العنوان الذي يبينه المحدد التالي "http://purl.org/dc/elements/١,١". كذلك توصي مواصفة مكتب المملكة المتحدة

لمشابكة المكتبات والمعلومات بأن تمثل أسماء عناصر دبلن كور بالأحرف الصغيرة (أي هكذا: "dc:title" وليس "dc:Title") .

ولترميز دبلن كور المقيد ينبغي إضافة بيان الاسم المميز للتعريف المرجعي لمقيدات دبلن كور ، بحيث يظهر على النحو التالي: "xmlns:dcterms=.". كما توصي مواصفة مكتب المملكة المتحدة لمشابكة المكتبات والمعلومات بتمثيل مقيدات توضيح العنصر على أنها عناصر elements وليس خصائص[1] attributes ، بحيث يمثل مقيد التاريخ المسمى " created " : - على سبيل المثال - هكذا:

```
<dcerms:created>٢٠٠٢</dcerms:created>
```

وليس هكذا :

```
<dc:date type="created">٢٠٠٢</dc:date>.
```

وفي المقابل، ينبغي أن تمثل مقيدات خطة الترميز باستخدام خاصية " scheme "، وأن تمثل اللغة التي تظهر بها "القيمة value " باستخدام خاصية "lang" في صيغة لغة الترميز الموسعة ، هكذا:

```
<?xml version="١,٠"?>
<metadata
        Xmlns=http://myorg.org/myapp/
        Xmlns:xsi="http://www.w٣.org/٢٠٠١/XMLSchema-instance"
```

(١) ذلك لأنه من المعروف في سياق لغة الترميز الموسعة أن الخاصية تتبع العنصر ولا تساويه [المترجم].

```
Xsi:schemaLocation="http:myorg.org/myapp/http://
http:myorg.org/myapp/scem.xsd"
xmlns:dc="http://purl.org/dc/elements/١،١/"
xmlns:dcterms="http://purl.org/dc/terms/"
<dc:title xml:lang="en">the electronic text center
introduction to TEI and guide to document
prepation</dc:title>
<dc:creator>Seaman, David</dc:creator>
<dcterms:created>١٩٩٥</dcterms:created>
<dc:identifier
schem="URL">http://etext.lib.verginia.edu/tei/
uvatei.html</cd:identifier>
</metadata>
```

ويمكن كذلك تمثيل دبلن كور في صيغة لغة الترميز الموسعة وفقاً لقواعد إطار وصف المصدر ، وقد أقرت مبادرة دبلن كور في أكتوبر ٢٠٠٢م، توصية مفادها " التعبير عن دبلن كور المبسط في صيغة لغة الترميز الموسعة القائمة على إطار وصف المصدر RDF/XML [(١)]. وبناء على هذه التوصية ،وفي حال استخدام إطار وصف المصدر ، ينبغي أن يبين ذلك عن طريق التاج المعبر عن هذا الإطار ، وهو التاج <rdf:RDF> ، مع ملاحظة إمكانية استخدام ترميز واحد لإطار وصف المصدر لتمثيل عدة مصادر طالما أن وصف كل مصدر من هذه المصادر سيرد محصوراً داخل عنصر وصف مستقل (<rdf:Description>) من عناصر وصف

(١) (راجع الرابط http://www.dublincore.org/documents/٢٠٠٢/٠٧/٣١/dcmes-xml/) [المترجم].

المصدر ، لكن لا يمكن استخدام أي مقيدات أو عناصر وصف معرفة داخلياً في هذا الصدد ، كما يلاحظ عدم إمكانية إدراج صيغة لغة الترميز الموسعة القائمة على إطار وصف المصدر الناتجة عن ذلك الاتجاه في صفحات الوب.

ويبدو الترميز الفعلي لعناصر دبلن كور واضحاً ومباشراً، كما يعكسه المثال التالي:

```
<?xml version="١,٠"?>
<!DOCTYPE rdf:RDF PUBLIC"..//DUBLIN CORE//DCMES DTD
٢٠٠٢/٠٧/٣١/EN/""http://Dublincore.org/document/٢٠٠٢/٠٧/٣١/dcesx
ml/dcmes-xml-dtd.dtd"/
<rdf:RDF xmlns:drf="http://www.w٣.org/١٩٩٩/٠٢/٢٢-drf-syntax-ns#"
 xmlns:dc=http://purl.org/dc/elements/١,١/>
<rdf;Description rdfabout=" http://etext.lib.verginia.edu/tei/
uvatei.html ">
<dc:title>the electronic text center introduction to TEI and guide to
document prepation</dc:title>
<dc:creator>Seaman, David</dc:creator>
  <dc:date>٢٠٠٢-٠٧-٣١</dc:date>
</rdf:Description>
</rdf:RDF>
```

مع ملاحظة أنه إذا كان للمصدر محدد واحد من المحددات الموحدة ، فيرمز على أنه قيمة للخاصية المسماة " rdf:about " ، وليس على أنه قيمة

لعنـصر المعـرف المـسمى <dc:identifier> ، أمـا إذا كـان لهـذا المـصدر محددات متعددة ، فيرد المحدد الإضافي أو المحددات الإضافية في عنصر المعرف(١).

وتجدر الإشارة إلى أنه ليست هنالك توصية معتمدة للتعبير عن دبلن كور المقيد في صيغة لغة الترميز الموسعة القائمة على إطار وصف المصدر ، وما هنالك سوى توصية مقترحة لمّا تعتمد بعد(٢). وبناءً على المستند الذي يتضمن هذه التوصية يمكن القول إن كلاً من عناصر دبلن كور ومقيدات توضيح هذه العناصر، تتوافقان مع الخصائص الأساسية والثانوية لإطار وصف المصدر، كما أن مقيدات خطة الترميز [في دبلن كور] ـ من ناحية أخرى - تتوافق مع "فئات classes"، أو " أنواع types" إطار وصف المصدر.

وفي المثال التالي الذي أعد وفقاً للتوصية المقترحة، ثمة تأكيدان: أولهما يرى أن قيمة مقيد توضيح العنصر المسمى " created " مشار إلى أنها ١٩٩٥م، وثانيهما أن الخاصية " created " نفسها ، يلاحظ أنها جاء كخاصية ثانوية لعنصر التاريخ في دبلن كور، هكذا :

```
<rdf:Description>
<dcq:created>١٩٩٥</dcq:created>
```

(١) أي المحددات التي لم تعامل على أنها المحدد الأساسي للمصدر [المترجم] .

(٢) Stefan Kokkelink and Roland Schwanzl, "Expressing Qualified Dublin Core in RDFI XML," ٢٩ August ٢٠٠١, available at

http://ww\dublincore.org/documents/٢٠٠١/٠٨/٢٩/dcq-rd-xml/. Accessed ٢٤ June ٢٠٠٢.

</rdf:Description>

</rdf:Description about="htpp://purl.org/dc/terms/created">

</rdfs:subPropertyOF

rdfresource="htpp://purl.org/dc/elements/١،١/date"/>

</rdf:Description>

وهناك متطلب تركيبي آخر غير إمكانية تمثيل دبلن كور البسيط والمقيد، وهو إيجاد نوع من الآلية للجمع بين عناصر من دبلن كور ومجموعات أخرى من عناصر ما وراء بيانات معرفة، حيث أدرك المنفذون منذ البداية أن الارتقاء بدبلن كور للاستخدام في مجالات تطبيقية معينة يستلزم أن تلحق به عناصر إضافية أخرى، ولعل من العناصر التي أثبتت جدواها للقيام بتلك المهمة عنصر "البادئة prefix "، حيث يشير إلى الخطة التي أخذ منها عنصر ما وراء البيانات.

ومن الأمثلة الجيدة على ذلك ما قدمه مشروع بيب لنك[١] BIBLINK الذي مولته المفوضية الأوربية ، فقد أعد هذا المشروع لتشجيع الناشرين على تقاسم مجموعة من ما وراء البيانات الوصفية المقننة لتستخدم في وصف الوثائق الإلكترونية من قبل الخدمات الببليوجرافية الوطنية والتي ستقوم بدورها بإعادة إرسال ما وراء بيانات محسنة إلى هؤلاء الناشرين.

---

(١) بيب لينك : مشروع مولته المفوضية الأوربية بهدف إنشاء علاقة بين الهيئات الببليوجرافية الوطنية وناشري المواد الإلكترونية [المترجم].

وبالفعل قام هذا المشروع بتعريف مجموعة ما وراء بيانات قوامها تسعة عشر عنصرًا[١]، استعير اثنا عشر عنصرًا منها من دبلن كور ، وخصصت سبعة منها لخدمة أهداف المشروع تحديداً ، من هذه العناصر : مجموع التدقيق[٢] checksum ، ومكان النشر، وتتابع الصدور، والسعر ، بحيث يستخدمها الوصف المتوافق مع بيب لينك في صيغة لغة ترميز النص الفائق كبادئات لتمييز عناصر دبلن كور المقننة عن العناصر الخاصة بمشروع بيب لينك، وذلك على النحو التالي:

```
<meta name="BIBLINK.checksum"
Content="fd٦٦e٣٧fb٦٩٣٤٩١le١٨٤b٠٩٢١٢٦٥">
<meta name="DC.Title" content=Taylor-Schechter United Home Page">
```

كما تجدر الإشارة إلى أن طريقة الاسم المميز namespace تقدم آلية رسمية لإمكانية التوسعة في صيغتي لغة الترميز الموسعة وإطار وصف المصدر. (لكن مع ملاحظة أن معرف نوع وثيقة لغة الترميز الموسعة XML DTD ، لا تدعم الأسماء المميزة بشكل صريح ، وبالتالي يفضل استخدام خطة لغة الترميز الموسعة لتعريف الوثيقة). ويعرض المثال التالي المأخوذ من " إرشادات تنفيذ دبلن كور بصيغة لغة الترميز الموسعة التي

---

(٢) Dublin Core qualifiers, available at http://dublincore.org/documents/dcmes-qualifiers/. Accessed ٢١ June ٢٠٠٢.

(١) قيمة يتم حسابها بناء على محتويات كتلة من البيانات وذلك لغرض الكشف عن صحة البيانات من فسادها [المترجم].

وضعها مكتب المملكة المتحدة لمشابكة المكتبات والمعلومات " تسجيلة تشتمل على عناصر دبلن كور وعنصر نموذج الكيان التعليمي(١) Learning Object Model المسمى " TypicalLearningTime (٢) لمعهد مهندسي الكهرباء والإلكترونيات IEEE :

```
<?xml version="١,٠"?>
<record
        xmlns="http://myorg.org/learningapp/"
        xmlns:xsi="http://www.w٣.org/٢٠٠١/XMLSchema-instance"
        xsi:schemaLocation="http://myorg.org/learningapp/http://myorg.
        org/learningapp/schema.xsd"
        xmlns:dc:="http://purl.ord/dc/elements/١,١/"
        xmlns:dc="http://www.imsglobal.org/xsd/imsmd_v١p٢">
        <dc:title>frog math<dc:title/>
        <dc:identifier>http://somewhere.com/frogmath/</dc:identifier>
        <dc:descriprion>simple math gamea for ٥-٧ year
olds.</dc:description>
        <ims:typicallearningtime>
                <ims:datetime>٠٠٠٠-٠٠-٠٠T٠٠:١٥</ims:datetime>
        </ims:typicallearningtime>
```

(٢) وردت في النص هكذا مع أن هذا الاختصار يشير إلى Learning Object Metadata [المترجم].

(٣) تعني " الوقت المثالي للتعليم" [المترجم].

</record>

الملفات التفصيلية أو بروفايلات التطبيق **application profiles** :

كما تبين من مثال بيب لنك الذي أشير إليه في الفقرة السابقة ، من المعتاد عند استخدام دبلن كور لوصف المصادر الخاصة بمشروع أو تطبيق معين أن يقوم المنفذون بإلحاقها عناصر أو مقيدات إضافية تفرضها طبيعة ذلك التطبيق ، بل قد يستشعر هؤلاء المنفذون الحاجة إلى وجود ضوابط صارمة تتصل بإجراءات استخدام تلك العناصر (كتعيين بعض العناصر الإجبارية على سبيل المثال) ، أو الحاجة إلى وجود إرشادات أكثر دقة تتصل بالمحتوى أكثر من تلك الإرشادات التي تضمنتها خطة دبلن كور نفسها.

وتعد بروفايلات التطبيق إحدى طرق إضفاء الطابع الرسمي على تعريف خطط ما وراء البيانات المعتمدة على دبلن كور، كما أن بروفايل التطبيق يعد من الناحية الرسمية خطة معدة لتطبيق معين ، ويتكون هذا البروفايل من عناصر بيانات مستمدة من خطة واحدة مسبقة التعريف أو أكثر من خطة واحدة، وتقتصر مهمة البروفايل على توضيح معاني العناصر التي تتضمنها الخطة، لكنه لا يتعداها إلى توسعة معانيها أو إضافة عناصر جديدة إلى تلك الخطة، وإن كان يمكنه أن يحدد ضوابط الاستخدام (كأن يبين مثلاً إذا ما كان العنصر يعامل على أنه عنصر إجباري ، أو عنصر غير متكرر) ، أو تحديد كيفية تمثيل البيانات المسموح باستخدامها أو البيانات الإجبارية ، أو تحديد اللغات المقيدة المعتمدة عند التطبيق .

ويفضل أن تنفذ بروفايلات التطبيق كخطة بصيغة لغة الترميز الموسعة ، حيث تدعم الأسماء المميزة ، إضافة إلى أن خطط الترميز الموسعة تدعم قيود الاستخدام المحلي ، كاستخدام القوائم الاستنادية لقيم العناصر ، وللعناصر الإجبارية ، كما أنها تدعم قيود مدى قابلية العناصر للتكرار. ويمكن كذلك تنفيذ بروفايلات التطبيق في صيغة إطار وصف المصدر ، ولكن بقدر أقل من المرونة . وعلاوة على ذلك يمكن نظرياً أن تنشأ بروفايلات التطبيق كاتفاقيات مكتوبة للمنفذين ومرمزة في أي تركيبة طالما أن قابلية فهمها آلياً وتنفيذها فنياً ليسا أمراً إلزامياً . وفي هذا السياق اتسع مفهوم فكرة بروفايلات التطبيق في بعض المجتمعات ليشمل تحديد نوع المعلومات التي ستظهر عادة في أدلة الاستخدام ، بما في ذلك اشتمالها على المزيد من إرشادات كيفية اختيار المحتوى وكيفية صياغته وبشكل يزيد عما تفرضه لغة خطة الترميز الموسعة .

وتعد خطة مشروع بيب لنك التي سبقت الإشارة إليها نموذجاً لبروفايلات التطبيق التي اختارت أن تمثل كخطة بلغة الترميز الموسعة حتى تحقق أهدافها(١)، كذلك ترعى مبادرة دبلن كور خطط تطوير بعض بروفايلات التطبيق مثل بروفايل التطبيقات المكتبية Libraries application profile الذي تتولى تطويره مجموعة عمل المكتبات بهدف دعم تطبيقات المكتبات لدبلن كور ، كاستخدامها له كشكل من أشكال أو صيغ التبادل

(١) Jane Hunter, An XML Schema Approach to Application Profiles, ٣ October ٢٠٠٠, available at http://archive.dstc.edu.au/maenad/appln_profiles.html. Accessed ٢٤ June ٢٠٠٢.

البينـي بـين الـنظم المـستخدمة لمعايير مـا وراء بـيانات مختلفة ، واستخدامه كذلك في تطبيقات جمع ما وراء البيانات ، كتلك التطبيقات التي تعتمد على بـروتوكول مبادرة الأرشيفات المفتوحة لجمـع مـا وراء بيانات، واستخدامه كذلك فـي إنشاء تسجيلات الفهرس المبسطة مع دلالات دبلن كور ، يضاف إلـى ذلك أنه لا تزال هناك بروفايلات تطبيق أخرى ترعاه مبادرة دبلن كور قـيد التطويـر، مـنها مـا يتـصل بالمجـالات الحكومـية ، ومـنها مـا يتـصل بالمجالات التعليمية والبيئية.

استخدامات وقضايا :

لقد ثبتت جدوى استخدام دبلـن كـور فـي العديد من المجالات المكتبية ، سواء اقترن ذلك بإضافة المصطلحات المتخصصة في أي من تلك المجالات التي تستخدم فيها ، أو من غير إضافة تلك المصطلحات ، فكثيراً ما يستخدم دبلـن كـور فـي البـوابات المتخصصة حيث يستنبط وصف لأي من المصادر التـي تظهر علـى صفحات تلك البـوابات مـن قاعدة بيانات مختصرة خاصة بمعلـومات دبلـن كـور. ومـن الـشائع كـذلك اسـتخدامه لوصـف الـنصوص الإلكترونية والـصور التـي تنشأ فـي سياق مـشروعات الـرقمنة الـراجعة ، وخاصة تلك المشروعات التي تضم عدداً كبيراً من المواد التي يصعب القيام بالفهرسـة الكاملـة لهـا أو لـيس هنالك مـا يسوغ ذلـك ، كما يستخدم كذلك مع بعض المـواد كالصور ومقالات الصحف التي يكتنف تطبيق القواعد الأنجلو أمـريكية علـى وصـفها شـيء مـن الـصعوبة ، أو يحـدث تطبيقها نـوعاً مـن الإشكالات التـي لا تشجع علـى تضمينها الفهرس العـام للمكتبة. وفـي هذه السياقات يحـاول مـستخدمو الخطـط المعتمدة علـى دبلـن كـور استثمار مزايا

التقييس ، بينما يسمح لمصممي مشروعات التطبيق بقدر ما من التفاوت في تعريف عناصر البيانات ووضع الإرشادات التي تعين على استخدامها، مثال ذلك أن يقرر في سياق أحد المشروعات استخدام القواعد الأنجلو أمريكية للفهرسة ، وما يتصل بها كالملفات الاستنادية للأسماء عند التعامل مع وصف عنصري "المنشئ" و"المشارك" ، ولكن لا تستخدم تلك القواعد عند صياغة عناوين المواد ، أو غير ذلك من الجوانب الأخرى للوصف الببليوجرافي.

وهنالك موقف آخر كثيراً ما يحدث عندما تختزن ما وراء البيانات في قاعدة البيانات المحلية اعتماداً على إحدى الخطط الأكثر اتساعاً من خطة دبلن كور ، إلا أنه ينبغي أن تحول إلى دبلن كور لتستخدم في سياق أحد الفهارس الموحدة ، أو في كشاف أحد محركات بحث الإنترنت ، أو في قاعدة بيانات خارجية أخرى تضم تسجيلات من مصادر متعددة ، وفي هذه الحالة تعمل دبلن كور على الأقل مرشدًا أو موجهًا denominator يمكن أن تربط بها عدة خطط معقدة بحيث يمكن بحث مجموعة متجانسة من عناصر البيانات ، ولعل من أفضل الأمثلة على ذلك بروتوكول مبادرة الأرشيفات المفتوحة لجمع ما وراء البيانات ، الذي يطلب - كحد أدنى - أن يكون لدى جميع المواقع المشتركة معه القدرة على تصدير دبلن كور غير المقيد.

وتجد الإشارة إلى أن الأهمية الكبيرة التي اكتسبتها دبلن كور لا تكمن في كونها خطة لوصف مصادر المعلومات وحسب ، وإنما في الدور الذي نهضت به مبادرة دبلن كور لما وراء البيانات الذي تمثل في تجميع العديد من الأطراف المتفرقة ذات الاهتمام المشترك ، كما أنها استحدثت منظمة حق لها أن تصبح دولية من حيث مجال نشاطها واتساع أبعاد المشاركة فيها، كما طرحت تلك

المبادرة على جبهة البحث كثيراً من القضايا المتصلة بلغة التمثيل، والتمثيل متعدد اللغات ، سواء [لعناصر] ما وراء البيانات أو لخطط ما وراء البيانات. يضاف إلى ذلك أن ورش العمل التي رعتها تلك المبادرة أوجدت أرضية مشتركة لتلاقح المعلومات وتبادلها بين المكتبات والمتاحف وغيرها من مؤسسات التراث الحضاري و الجهات الحكومية والهيئات العلمية ومطوري الوب وعلماء الحاسبات والمعلمين وغيرهم ، مما كان له الأثر الإيجابي في الارتقاء بتلك المجتمعات كافة .

وأدَّت مبادرة دبلن كور لما وراء البيانات كذلك دوراً مهماً في إحاطتها مجتمع المكتبات بقضايا التشغيل البيني interoperability الخارج عن إطار بيئة مارك المغلقة؛ فمنذ البدايات المبكرة لتطوير البنية الفنية النظرية التي تكفل تجميع ما وراء البيانات من الخطط المتنوعة (كما كان الحال مع "إطار وارويك warwick") إلى الاستخدام الحالي للأسماء المميزة في صيغة لغة الترميز الموسعة ، وصيغة إطار وصف المصدر ، يقر الباحثون المرتبطون بدبلن كور على الدوام بحاجة العالم الحقيقية إلى ضرورة التكامل بين ما وراء البيانات الوصفية وما وراء البيانات الإدارية المأخوذة من مصادر مختلفة في أوقات مختلفة. كما طورت في السياق نفسه نقاط المقابلة أو التحويل البيني crosswalks من جميع خطط ما وراء البيانات الوصفية المهمة تقريباً إلى دبلن كور. كذلك حاولت عدة مشروعات أولية prototype كانت قد طورت لدعم جهود مكتب تسجيل مبادرة دبلن كور DCMI registry الجمع بين المصطلحات المأخوذة من بروفايلات تطبيقات وخطط خاصة بمجالات ذات صلة بها ، وكذلك من الأسماء المميزة الرسمية لدبلن

كور. وفضلاً عن ذلك فقد كان للاشتراك المكثف لكثير من أعضاء مبادرة دبلن كور في مجريات جهود تطوير إطار وصف المصدر والوب الدلالية أثره الإيجابي في الارتقاء بوعي مجتمع المكتبات بهذه المبادرات. ومن الملاحظ كذلك أن مبادرة دبلن كور سعت في السنوات الأخيرة لتوسعة مجال رسالتها عن قصد لتصبح منتدى عاماً لمناقشة قضايا الاستكشاف متداخل المجالات cross-domain discovery ، وقضايا أطر العمل الخاصة بالتشغيل البيني .

وفي الوقت الذي تبرز فيه تلك الجهود ، توجه فيه انتقادات لمبادرة دبلن كور لما وراء البيانات على تأخرها في إصدار الإرشادات الأساسية لمنفذي دبلن كور ، وانتقادات أخرى لخلو الساحة حتى الآن من توصيات معتمدة بشأن التمثيل التركيبي لدبلن كور المقيد في صيغتي لغة الترميز الموسعة ، وإطار وصف المصدر، وإن كانت هناك إرشادات قيد التطوير منذ عام ١٩٩٨م لتمثيل الإرجاعات الببليوجرافية بمقالات الدوريات في دبلن كور، لكنها لما تنته بعد .

كما تجدر الإشارة في هذا الصدد إلى أن مقيدات توضيح عناصر البيانات الخاصة بكل من "المنشئ" ، و"المشارك" ، و"الناشر" قد ألغيت من [قائمة] مقيدات دبلن كور ، لا لشيء إلا لعدم انتهاء أعضاء لجنة الاستخدام إلى إجماع بشأنها ، ولا يزال القرار بشأنها معلقاً بالرغم من إبداء بعض المنفذين حاجتهم الملحة لها ، وما زلنا في انتظار إذا ما كانت احتياجات المنفذين ستلبى أم ستظل معلقة على هذا النحو ، ولعل تفسيرنا لذلك أن مبادرة دبلن كور لما وراء البيانات تولي اهتماماً متزايداً بالقضايا النظرية

والتطبيقية المتصلة بعملية التشغيل البيني أكثر من غيرها من القضايا الأخرى .

READINGS :

- Dekkers, Makx, and Stuart L. Weibel. "Dublin Core Metadata Initiative Progress Report and Workplan for ٢٠٠٢." D-Lib Magazine ٨, no. ٢ (February ٢٠٠٢). Available at http://www.dlib.org/dlib februaryO٢/weibeI/٠٢weibel.html.

  A "state of Dublin Core" report from ٢٠٠٢. D-Lib Magazine has published numerous articles on Dublin Core, including reports from most of the Dublin Core workshops. These reports, read in order, constitute a history of the evolution of the DCMI.

- Dublin Core Metadata Initiative (home page). Available at http://www.dublincore.org/.

  Many relevant documents are linked to from the Dublin Core home page, including the current reference description of the Dublin Core specification itself

  (http://www.dublincore.org/documents/dces/) and the Dublin Core qualifiers (http://www.dublincore.org/documents/dcmes-qualifiers/.) Approved DCMI Recommendations and proposed Recommendations are listed under the link "Documents." The "Resources" link includes a bibliography of writings related to Dublin Core.

- Guenther, Rebecca, and Priscilla Caplan. "Metadata for Internet Resources: The Dublin Core Metadata Elements Set and Its Mapping to USMARC." Cataloging and Classification Quarterly ٢٢, ٠,٣/٤ (١٩٩٦).

  A discussion of some of the issues raised in attempting to map from Dublin Core to MARC. The Library of Congress maintains an official crosswalk from MARC to Dublin Core at http://lcweb.loc.gov/marc/marc٢dc.html and between Dublin Core, MARC٢١ , and GILS at http://www.loc.gov/marc/dccross.html.

الفصل التاسع

الوصف الأرشيفي و الوصف الأرشيفي المرمز

تعرف المجموعة الأرشيفية (أو الأرشيفات) بأنها "مجموعة منظمة من السجلات غير الجارية الخاصة بإحدى المؤسسات أو الجهات الحكومية أو المنظمات أو الهيئات ، أو أنها الأوراق الشخصية لأحد الأفراد أو إحدى العائلات، والتي يحتفظ بها في مستودع معين نظراً لقيمتها التاريخية(١) ". وعادة ما تضطلع بمسؤولية هذه المجموعة الأرشيفية عدة هيئات ، مثل الهيئات الحكومية ، والوحدات الخاصة بهذه المجموعات في المؤسسات الربحية أو غير الربحية ، هذا فضلاً عن الأقسام المعنية بالمجموعات الخاصة والمخطوطات بالمكتبات.

مبادئ الوصف الأرشيفي :

إذا كان لمهنة المكتبات - كما أشرنا في الفصل السادس - تقليدها المديد في جهود الوصف الببليوجرافي الذي تعكسه في الوقت الحالي الإصدارة المنقحة من القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية AACR٢R ، فقد كان للأرشيفيين تقليدهم كذلك في الوصف الأرشيفي الذي يختلف عن الوصف الببليوجرافي في عدة أوجه ؛ فبينما يركز الوصف الببليوجرافي على المطبوع ويعامله كوحدة مستقلة بذاتها ، يعنى الوصف الأرشيفي بالتجميعات aggregations التي تعرف عند الإشارة إلى الأوراق الخاصة

(١) ODLIS: Online Dictionary of Library and Information Science, available at http://vax.wcsu.edu/library/odlis.html. Accessed ٢٤ June ٢٠٠٢.

بإحدى الجهات بمجموعة المحفوظات(١) record group ، وعند الإشارة إلى الأوراق الشخصية الخاصة بأحد الأفراد بالمجموعة الأرشيفية archival collection.(مع ملاحظة أن كلتيهما تسمى في عرف الأرشيفات الأنجلو - كندية بالوحدة الأرشيفية الرئيسية أو المتكاملة(٢) fonds).

وقد جرت العادة أن تجمع مواد هذه التجميعات بعضها مع بعضٍ، إما على أساس الأصل أو المنشأ provenance (٣)، أو على أساس تاريخ إنشائها أو ملكيتها. وبحكم أن الوحدة الأساسية للوصف الأرشيفي هي التجميعةِ aggregation ، فلا يحظى وصف السمات المادية للمواد التي تشكل المجموعة بالقدر نفسه الذي يحظى به وصف السمات الفكرية ، بل قد ينظر إلى ذلك الجانب الأول على أنه أقل أهمية من ذلك الجانب الأخير.

---

(٢) مجموعة من المحفوظات المتعلقة بمؤسسة واحدة والتي تكونت على أساس مبدأ المنشأ أو الملكية provenance بواسطة هيئة أرشيفية archives بغرض ضبطها ، وتلك المجموعة تؤلف محفوظات هيئة أو مؤسسة أرشيفية ذات كيان مستقل [المترجم] .

(١) مصطلح شاع في القارة الأوربية وإلى حد ما في شمال القارة الأمريكية يشير إلى الوحدة الأرشيفية الرئيسية التي هي أساس جميع قواعد ترتيب محتويات الأرشيف . ويشير المصطلح إلى مجموعة الوثائق التي تم إنشاؤها ونمت نمواً طبيعياً نتيجة للأنشطة والمعاملات الإدارية والتي تعتبر كلاً متكاملاً كما أنها تصلح لأن تساند العمل بالمؤسسة بدون أية مساعدات من مجموعة وثائق أخرى ،وتسمى كذلك بالمجموعة الأرشيفية [المترجم].

(٢) تعني عدة أشياء في مجال الأرشيف منها المصدر أو الأصل أو مكتب المنشأ ، أي الكيان الإداري المسؤول عن نشأة المخطوطات أو الذي تسلمها وجمعها أثناء تشغيل أعماله ، وكذلك يشير إلى الأفراد والعائلات والمؤسسات أو أي مصدر مسؤول عن الأوراق الشخصية ومجموعات المخطوطات ، كما يشير إلى المصدر الذي أنشأ الوثيقة أو مصدر الأوراق الشخصية أو الكيان المنشئ أو الذي كون مجموعة المخطوطات [المترجم].

وثمة قاعدتان قويتان يقوم عليهما التوثيق الأرشيفي هما: احترام الوحدة الأرشيفية المتكاملة respect des fonds ، ومبدأ الترتيب الأصلي principle of original order ؛ حيث تقر أولاهما ، والتي تعرف كذلك بمبدأ المنشأ أو الأصل principle of provenance ، بضرورة حفظ المواد ذات الأصل الواحد بعضها مع بعض وعدم الخلط بينها وبين المواد الأخرى(١). أما ثانيتهما فتنادي بالحفاظ على الترتيب الأصلي الذي وضعت فيه المجموعة عند إنشائها لأول مرة(٢) . إلا أنه قد تكون هناك عند التعامل مع بعض المحفوظات الخاصة بمؤسسة معينة بعض المتطلبات أو الضوابط القانونية الخاصة بحفظها retention (٣)، والحفاظ على تكاملها(٤) integrity ، وإثبات ملكيتها(٥) authenticity . ومع ذلك يسعى الوصف الأرشيفي إلى التأكيد على أن لتوثيق كل من أصل المجموعة ،وترتيبها الأصلي الأهمية القصوى التي لا تضارعهما فيه أهمية أخرى.

---

(١) في النظرية الأرشيفية ، المبدأ الذي ينادي بأن المحفوظات أو الأرشيف التابع لمنشأ أو مصدر معين لا يجب أن يخلط مع الأرشيفات الأخرى أو المحفوظات الأخرى التابعة لمصادر أخرى . [المترجم].

(٢) مبدأ ينادي بأن تحفظ الوثائق في شكلها أو بنيتها الأصلية ، وكذلك يجب أن تبقى في ترتيبها الأصلي ، وذلك حتى يمكن الحفاظ على علاقات الوثائق بعضها ببعض في وحدة متكاملة ومنظمة [المترجم].

(٣) يقصد بها مدة حفظ الوثائق قبل ترحيلها إلى الأرشيف [المترجم].

(٤) مبدأ من مبادئ تنظيم الأرشيف ينص على أن المجموعة الأرشيفية المتكاملة يجب أن تحفظ بدون تقسيم [المترجم].

(٥) الوثائق موثوق فيها فيما يتعلق بمنشئها [المترجم].

وعادة ما يبدأ الوصف الأرشيفي بوصف مجموعة المحفوظات أو المجموعة الأرشيفية وصفاً عاماً ينظر فيه إلى هذه المحفوظات على أنها مجموعة متكاملة، ثم يتدرج في وصف الحلقات أو السلاسل المختلفة ، ثم السلاسل الفرعية التابعة لها ، وذلك في أسلوب هرمي أو طبقي يستمر الوصف فيه حتى يصل للمستوى الأدنى الذي توجد فيه المواد الفردية ، وإن كان من الممكن أن يقف الوصف عند مستوى أعلى من ذلك المستوى الأدنى.

وعندما نشرت القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية عام ١٩٧٨م كانت تتضمن فصلاً مستقلاً عن فهرسة المخطوطات ، لكن رأى الأرشيفيون على مستوى العالم أنه لا يصلح للتطبيق على المجموعات الأرشيفية والمخطوطات لعدم مراعاته المبادئ الراسخة للوصف الأرشيفي، حيث عني بوصف المواد الفردية على حساب التجميعات ، كما ركز على الوصف المادي أكثر من تركيزه على توثيق الأصل أو المنشأ، مما دفع ستيفين هينسن Steven Hensen من مكتبة الكونجرس إلى إعداد مسودة الدليل الإرشادي لفهرسة المستودعات الأرشيفية المسمى " الأرشيفات ، والأوراق الشخصية، والمخطوطات Archives, Personal Papers, and Manuscripts (APPM)" ، والتي أصبحت فور نشرها في عام ١٩٨٣م معياراً لفهرسة مجموعات الأرشيفات والمخطوطات(١).

(١) Steven L. Hensen, Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Libraries, ٢nd ed. (Washington, D.C.: Society of American Archivists, [١٩٨٣], ١٩٨٩).

وفي الفترة نفسها تقريباً سعت إحدى المجموعات التي شكلتها جمعية الأرشيفيين الأمريكيين Society of American Archivists(SAA) بتمويل من "الهبات الوطنية للإنسانيات " والتي تعرف بلجنة المهام لنظم المعلومات الوطنية the National Information System Task force ، إلى تطوير صيغة مارك الأمريكي لضبط المجموعات الأرشيفية والمخطوطات USMARC format for Archives and Manuscripts Control (AMC) ، ورأى الأرشيفيون وأمناء المخطوطات استخدام ذلك الدليل الإرشادي كمعيار محتوى عند إنشاء التسجيلات وفق هذه الصيغة المقترحة ، كما أصبح فهرس ونظام الفهرسة الخاص بمجموعة مكتبات البحث التابعة لشبكة معلومات المكتبات البحثية[١] (RLIN) هو الفهرس الموحد المباشر الأساسي لمجموعات الأرشيفات والمخطوطات في الولايات المتحدة.

ويمكن القول إن نظام الفهرسة الخاص بصيغة مارك الأمريكي لضبط المجموعات الأرشيفية قدم مستوى لم يسبق له مثيل من حيث إمكانات الوصول إلى المجموعات الأرشيفية المهمة ، إلا أن هذه الصيغة وبسبب القيود التي تفرض على طول تسجيلات مارك وبنيتها لم تستطع أن تقوم مقام تلك الأدوات الأكثر تفصيلاً في هذا الصدد والمسماة "معينات البحث finding aids" التي استخدمتها المستودعات الأرشيفية منذ فترات بعيدة .

(١) مؤسسة غير ربحية من مجموعة المكتبات والأرشيفات والمتاحف وبعض مؤسسات التراث الثقافي الأخرى. مهمتها تحسين عمليات الوصول إلى المعلومات التي تساند عمليات البحث والتعليم ، و من الخدمات التي تقدمها RLIN قاعدة بيانات ببليوجرافية للمجموعات المكتبية والأرشيفية [المترجم].

وتعد معينة البحث هذه والتي تأتي عادة في شكل أداة الجرد[1] inventory ، أو السجل الدفتري[2] register ، أو التقويم الزمني[3] calendar ،هي الأداة الأساسية التي تستخدم لإجراء الضبط الإداري والفكري للمجموعات الأرشيفية والمخطوطات. ونتيجة لعدم توافر المعايير الرسمية لمحتوى معينات البحث حتى وقت قريب ، فإن أشكال هذه المعينات ومحتواها يتفاوت تفاوتاً كبيراً من مستودع لآخر ، بل من مجموعة أرشيفية لأخرى. وبصرف النظر عن ذلك يلاحظ أن هذه المعينات والتزامها بمبادئ الوصف الأرشيفي عادة ما تبدأ بوصف أصل المجموعة الأرشيفية أو مصدرها وصفاً تفصيلياً، بحيث يمكن أن يتطرق هذا الوصف إلى إعطاء سيرة ذاتية مختصرة لمالك المجموعة ، أو تاريخاً للمؤسسة التي تمتلكها ، أو تعريفاً بالجهة التي تقتنيها ، ذلك حسب طبيعة المجموعة التي يتم وصفها ، ثم يتطرق الوصف بعد ذلك إلى وصف مجال مواد المجموعة ومحتواها ، ثم وصف التجميعات الفردية للمواد (أي السلاسل و السلاسل الفرعية) ، ثم وصف الملفات (أو الحاويات container) ، وقد يستمر الوصف ليصل في بعض الحالات إلى مفردات هذه التجميعات. ويعرض الشكل رقم (١/٩)

---

(٢) تعني في الأرشيف وسيلة إيجاد للمواد في مجموعة أرشيفية مرتبة بترتيب المواد المحفوظة نفسها وتعني كذلك في تنظيم المحفوظات قائمة الوثائق التي يزمع التخلص منها[المترجم].

(٣) قائمة في شكل مجلد بالحوادث أو الخطابات الصادرة والواردة او الإجراءات التي تتخذ ، في ترتيب مبسط عادة حيث ترتب ترتيبًا زمنيًا بالرقم المسلسل وتستعمل وسيلة لإيجاد الوثائق [المترجم].

(٤) قائمة مرتبة تاريخيًا لمجموعة شاملة من الوثائق تعطي تاريخ الوثيقة ملخصًا عن محتوياتها [المترجم].

إحدى معينات البحث المختصرة إلى حد ما.

بنية الوصف الأرشيفي المرمز وعناصره :

طور الوصف الأرشيفي المرمز في تسعينات القرن العشرين الميلادي لترميز معينات البحث التقليدية المطبوعة في صيغة مقروءة آلياً، ولكن لعدم وجود معيار واحد يعتمد عليه على المستوى الدولي عند إنشاء معينات البحث - كما أشرنا من قبل - فقد قام منشئو الوصف الأرشيفي المرمز بجمع عينات من معينات البحث من عدد من المستودعات ، وحاولوا استيعاب نطاق الممارسات المختلفة ، ثم اختيرت لغة الترميز المعيارية العامة SGML كخطة ترميز هذه المعينات ، ذلك لما يتوافر لها من إمكانات على معالجة النصوص المكتوبة المطولة والمستويات المتعددة للتسلسل الطبقي ، وأصبح للوصف الأرشيفي المرمز الذي نفذ في أول الأمر كمعرف لنوع الوثيقة في صيغة لغة الترميز المعيارية العامة SGML DTD ، معرف آخر لنوع الوثيقة في صيغة لغة الترميز الموسعة XML DTD .

ويتكون معرف نوع الوثيقة الخاص بالوصف الأرشيفي المرمز EAD DTD من ثلاثة أقسام رئيسية ، هي:

- قسم الترويسة(1) <eadheader> ، يشتمل على المعلومات التي تخص الوصف الأرشيفي المرمز نفسه.

الشكل رقم (١/٩)

إحدى معينات البحث معروضة بشكل مختصر من بدايتها إلى الجزء الخاص بقائمة

(١) تسمية الأقسام بهذه الأسماء من اجتهاد المترجم ، وما ورد في الأصل هو ذكر مسمى التاج فقط، هكذا : <eadheader> وهلم جرًا [المترجم].

المحتويات

Summary Information

Title: W. May Walker Papers
Inclusive Dates: 1901-1974
Bulk Dates: 1925-1974
Call No.: MSS 75-12
Creator: Pansy Walker
Extent: 12.75 Linear/Cubic Feet; 53 Boxes
Repository:
Special Collections, Florida State University Libraries

Storage: Box 889-971 shelved at Claude Pepper Library, Florida State University. For current information on the location of materials, please consult the Special Collections Department's home page.
Abstract:
The collection includes correspondence, committee papers, legislative papers, and legal papers such as rulings, opinions, and papers concerning particular cases. Judge Walker served as County Judge of Leon County and on the bench of the 2nd Judicial Circuit. The great majority of the papers in the collection cover the time period from the 1940s until the time of his death, September 16, 1974. Also included among his papers are papers belonging to his father, Nat R. Walker, who was one of the most colorful and significant characters in the history of Florida.

Administrative Information

Acquisition Information:

The W. May Walker Papers were donated to the Special Collections Department, Florida State University Libraries, by Pansy Walker of Tallahassee, Florida, in 1976.

Access:

Preferred Citation:

[Identification of item], W. May Walker Papers, Special Collections, Florida State University Libraries, Tallahassee, Florida.

Biography

The Honorable W. May Walker was born in Crawfordville, Wakulla County, Florida on May 2, 1905, the son of Nat R. Walker and Alice (Tully) Walker, both of Crawfordville. After attending public schools in Leon County he received his Bachelor of Laws degree from Cumberland University, Lebanon, Tennessee, in 1927. That same year, after passing the Florida Bar exam, he began the practice of Law in Tallahassee.

Judge Walker then served as County Judge of Leon County from 1932 to 1940 when he took over the Circuit Court Judgeship after the death of J. B. Johnson. He served on the bench of the 2nd Judicial Circuit until the time of his death, September 16, 1974.

He married Pansy Crosby daughter of Joseph and Nora Horton Crosby in 1937 and they have two sons, W. May Walker, Jr. and Joseph Stanley Walker. During his lengthy career judge Walker has presided over many important and sensational cases and was one of the most hard-working and distinguished jurists of the State.

Collection Scope and Content Note

The papers of Judge W. May Walker cover all aspects of his legal career including correspondence, committee papers, legislative papers, rulings and opinions, as well as papers concerning particular cases. Judge Walker served as County Judge of Leon County and on the bench of the 2nd Judicial Circuit. The great majority of the papers in the collection cover the time period from the 1940s until the time of his death, September 16, 1974. Also included among his papers are papers belonging to his father, Nat R. Walker, who was one of the most colorful and significant characters in the history of Florida. A brief biography of Nat Walker is filed with his papers.

Controlled Access Terms

Note:

The following terms have been used to index the description of this collection in the Library's online catalog:

Subject Terms:

- Walker, W. May, 1905-1974
- Walker, Nat R.
- Judges -- Florida

Contents List

Container / Location Title

Series: Series A: Correspondence

- قسم الحالة الأساسية <frontmatter> ، ويعطي وصفاً يهدف إلى تهيئة معينة البحث لأغراض النشر.
- قسم الوصف الأرشيفي <archdesc> ، ويصف المجموعات الأرشيفية أو مجموعة المخطوطات.

ويشتمل قسم الترويسة <eadheader> على ثلاثة أقسام فرعية تأتي مشابهة إلى حد كبير لتلك التي تشتمل عليها ترويسة مبادرة ترميز النص TEI ، هي: وصف الملف <fileDesc> ، ووصف البروفايل <profiledesc> ، ووصف التنقيح <revisiondesc> . ويتضمن القسم الفرعي الخاص بوصف الملف <fileDesc> عناصر تتصل بعنوان معينة البحث وبيانات نشرها، أما القسم الفرعي لوصف البروفايل <profiledesc> فيتضمن تاريخ إنشاء معينة البحث، واللغة التي قدمت بها ، بينما يسجل القسم الفرعي لوصف المراجعة <revisiondesc> التغييرات التي أحدثت على معينة البحث مع مرور الوقت.

أما القسم الثاني - أي الحالة الأساسية <frontmatter> - فيشتمل على معلومات مشابهة لتلك التي يحتوي عليها القسم الفرعي الخاص بوصف الملف <filedsec> من قسم الترويسة <eadheader> وقد صيغ ليؤدي وظيفة صفحة العنوان المطبوعة لمعينة البحث ، وهو قسم نادر الاستخدام على أية حال.

أما قسم الوصف الأرشيفي <archdesc> فيعد بمثابة القلب بالنسبة للوصف الأرشيفي المرمز ، حيث يصف المجموعة الأرشيفية أو مجموعة المحفوظات نفسها وليس معينة البحث ، وفي سياق هذا القسم يمكن تكرار عناصر البيانات الوصفية التي تنطبق على أعلى مستوى (للمجموعة) على

كل وحدة فرعية بداخلها وعلى الوحدات الفرعية الأصغر التي تتضمنها تلك الوحدات الفرعية ، وبشكل يبدو فيه شكل الوصف الأرشيفي <archdesc> طبقيًا أو متواتراً أو مكرراً.

ونعرض فيما يلي تصوراً فكرياً يعرض الكيفية التي ترد بها العناصر الفرعية عالية المستوى داخل قسم الوصف الأرشيفي<archdesc>:

<did>descriptive identification
<admininfor>administrative information
<bioghist>biography or history
<scopecontent> scope and content
<organization> organization
<arrangement> arrangement
<note>note
<dao>digital archival object
<dagroup> digital archival object group
<controlaccess> controlled access headings
<add>adjunct descriptive data
<odd> other descriptive data
<dsc>description of subordinate components
<c٠١>component (١ st level)
<did>
<admininfor>
<bioghist>
<scopecontent>
<organization>
<arrangement>
<note>
<dao>

```
<dagroup>
<controlaccess>
<add>
<odd>
 <c٠٢>component (٢ nd level)
        <di>
        …
```

كذلك يشتمل عنصر الهوية أو التعريف الوصفي <did> على الوصف الأساسي للمواد في جميع مستوياتها. ومن ثم يشتمل بداخله على العناصر الفرعية التالية[١]:

- اسم المستودع الذي تقتنى بداخله المجموعة <repository>
- أصل المواد أو مصدرها <origination>
- عنوان الوحدة التي يتم وصفها <unittile>
- تواريخ إنشاء المواد التي تتضمنها الوحدة <unitdate>
- الوصف المادي للمواد <physdesc>
- الوصف المختصر للمواد <abstract>
- المعرف الذي يعطى للوحدة <unitid>
- المكان الفعلي الذي توضع فيه الوحدة <physloc>

ومع أن كل عنصر من هذه العناصر قد يقع في أي مستوى من مستويات الوصف، إلا أن بعضها (مثل اسم المستودع) يكون من الأنسب له أن يستخدم في المستوى الخاص بالمجموعة ؛ بينما يكون من الأنسب لبعضها الآخر (مثل الموقع) أن يستخدم في المستويات الدنيا.

---

(١) تسمية العناصر الفرعية بهذه الأسماء من اجتهاد المترجم ، وما ورد في الأصل هو ذكر مسمى التاج فقط ، هكذا : <repository> وهلم جرًا [المترجم].

وتجدر الإشارة إلى أن إدخال البيانات في معظم هذه العناصر الفرعية الكبرى لعنصر الهوية أو التعريف الوصفي <did> ، يمكن أن يتم بشكل مباشر ، أو بتمديد مجال ترميز العناصر الفرعية والخصائص. وفيما يلي ترميزٌ مبسطٌ لإحدى المجموعات - يعرضه الشكل رقم (٢/٩) - وترميزٌ مبسطٌ للمجموعة نفسها يعرضه الشكل (٣/٩).

```
<did>
   <repository>Harry Ransom Humanities Research Center</repository>
   <origination>Stoppard, Tom</origination>
   <unittitle>Tom Stoppard Papers</unittitle>
   <unitdate>1944-1995</unitdate>
   <physdesc>88 boxes (36 linear feet)</physdesc>
   <abstract>The papers of British playwright Tom Stoppard (b.
   1937) encompass his entire career and consist of multiple
   drafts of his plays, from the well-known
   <title render="italic">Rosencrantz and Guildenstern Are
   Dead</title> to several that were never produced,
   correspondence, photographs, and posters, as well as
   materials from stage, screen, and radio productions around
   the world.</abstract>
</did>
```

من

```
<did>
    <head>Summary Description of the Tom Stoppard Papers</head>
    <repository>
        <corpname>The University of Texas at Austin
        <subarea>Harry Ransom Humanities Research Center</subarea>
        </corpname>
    </repository>
    <origination>
        <persname source="lcnaf" encodinganalog="100">Stoppard,
        Tom</persname>
    </origination>
    <unittitle encodinganalog="245">Tom Stoppard Papers, </unittitle>
    <unitdate type="inclusive">1946-1995</unitdate>
    <physdesc encodinganalog="300">
        <extent>63 boxes (26 linear feet)</extent>
    </physdesc>
    <unitid type="accession">[illegible]</unitid>
    <physloc audience="internal">ldx:SW 5-3</physloc>
    <abstract>The papers of British playwright Tom Stoppard (b.
    1937) document his entire career and consist of multiple
    drafts of his plays, from the well-known
    <title render="italic">Rosencrantz and Guildenstern Are
    Dead</title> to several that were never produced,
    correspondence, photographs, and posters, as well as
    materials from stage, screen, and radio productions from
    around the world.</abstract>
</did>
```


كثر

E

أ ... علـى معلومات تتصل بإدارة المجموعة أو إحدى وحداتها الفرعية، مثل معلومات الاقتـناء والمعالجـة ، تـاريخ الوصـاية[1] custodial history ، وضـوابط الوصـول أو الاستخدام أو كليهما ، والصيغة المفضلة للإشارة المرجعية للمادة.

(١) تـاريخ الوصـاية يشير في مجال الوثائق والأرشيف إلى تتابع الإدارات أو الأشخاص الأمناء علـى المـواد الأرشـيفية منذ بـدء إنـشائها حتـى بدايـة الحـصول عليها بواسـطة الأرشـيف أو مخزن المخطوطات [المترجم] .

أما عنصر المعلومات الشخصية <bioghist> فيشتمل على نبذة عن سيرة الشخص أو عن تاريخ الهيئة في صيغة روائية أو تأريخية chronological. ويستخدم عنصر مجال المحتوى <scopecontent> لعمل ملخص للتغطية الموضوعية للمجموعة أو الوحدة الفرعية. أما عنصرا الترتيب <arrangement> والتنظيم <organization> الفرعيين فيمكن استخدامهما داخل عنصر مجال المحتوى لتحديد طبيعة هذه المعلومات ، كما يمكن لهذين العنصرين أن يأتيا خارجين عن هذا العنصر، أي مجال المحتوى على مستوى الوصف نفسه.

أما عنصر ضبط الإتاحة <controlaccess> فيعتبر عنصر غلاف wrapper element لترميز الصيغ المقيدة للأسماء والموضوعات التي قصد بها أن تكون نقاط إتاحة ، ومن ثم تشمل العناصر الفرعية كلاً من الأنواع المختلفة للأسماء (أي أسماء الهيئات <corpname> ، وأسماء الشهرة <famname> ، والأسماء الشخصية <persname>) ، وأسماء الأماكن (أي الأسماء الجغرافية <geogname>) ، وأسماء الموضوعات (<subject>) ، وأسماء الأشياء أو الأجناس (<genreform>) . ومع جميع هذه العناصر الفرعية يجب أن يحدد الملف الاستنادي المعتمد عليه في صياغة هذه الأسماء وذلك باستخدام خاصية "مصدر source ". هكذا :

```
<controlaccess>
<head>Subjects </head>
<subject source="lcsh">Civil War- Florida</subject>
<subject source="lcsh">Railroads- Florida</subject>
</controlaccess>
```

كذلك نلاحظ أن عنصر وصف المكونات التابعة <dsc> يعمل بمثابة غلاف لوصف الأجزاء التابعة أو المكونة ، كما أن هذه الأجزاء المكونة نفسها ترمز داخل عنصر شامل أو حاو container يمكن تمثيله عن طريق عملية ترقيم أو من دون ترقيم، أي أن ترد هذه الأجزاء مرقمة كعناصر مكررة باستخدام التاج <c> ، أو هكذا : <c٠n> ... ,<c٠٢> , <c٠١>، وإذا استخدم الترقيم فإنه يستمر فقط عند تداخل أجزاء المكونات بعضها داخل بعض حتى يتم بيان مستوى التداخل. هكذا:

```
<c٠١ level="series"><did><unittitle>series one</unittitle></did>
        <c٠٢   level="subseries"><did><unittitle>Subseries
        one</unittitle></did>
        <c٠٢   level="subseries"><did><unittitle>Subseries
        two</unittitle></did>
        </c٠٢>
</c٠١>
<c٠١   level="series"><did><unittitle>Series
two</did></unittitle></c٠١>
```

ويمكن لعنصر الحاوية <container> أن يبين التحديد المنطقي للصناديق boxes والمجلدات وغيرهما من الحاويات المادية الأخرى في أي مستوى. ونظراً لأن معظم المواد الأرشيفية لا تتوافر على الأرفف المفتوحة ليقوم المستفيد النهائي باسترجاعها ؛ فغالباً ما يلغى المكان المادي من الوصف . هكذا :

```
        <did>
                <container type="box">١-١٤</container>
                <unittitle>Campaign materials</unittitle>
```

....

وتتفاوت معينات البحث الأرشيفية في الكيفية التي تمثل بها المعلومات للمستفيدين؛ فبعض معينات البحث تجمع معلومات مختصرة عن الوحدات الفرعية الكبيرة مع بعض، مثل السلاسل ، ثم تحدد بعد ذلك في النهاية محتويات حاويات معينة ، بينما تقوم بعض المعينات الأخرى بوصف كل وحدة فرعية مع أحد سجلات الحاويات في الوحدة الفرعية. وتبين خاصية " النوع type " للعنصر الغلاف المسمى "[١] <dsc> " التركيبية المتبعة في الوصف الأرشيفي المرمز ، وذلك من خلال القيمة "نظرة عامة تحليلية analyticoverview" ، التي تبين الصيغة الأولى ، والقيمة "تجميعي combined" ، التي تبين الصيغة الثانية.

كذلك يدعم الوصف الأرشيفي المرمز الربط الداخلي من أحد أجزاء معينة البحث لجزء آخر منها، وكذا الربط الخارجي للملفات الأخرى. وقد تم تطوير بعض عناصر ربط الوصف الأرشيفي المرمز لتدعم مواصفتي الرابط الموسع في صيغة لغة الترميز الموسعة XML XLink ، والمؤشر الموسع XPointer ، توقعاً للدعم الواسع لتلك الخصائص، مع دعمها كذلك لعملية الربط المشابهة للربط في صيغة لغة ترميز النص الفائق HTML-like linking الأبسط منها. ومن الملاحظ أن جميع العناصر المحددة تدعم الخاصية المسماة "id" التي يمكن استخدامها لتحديد العنصر كهدف لرابط داخلي. ويمكن تأسيس الروابط الخارجية للوثائق ذات الصلة مثل معينات البحث الأخرى، أو للتمثيلات الرقمية للكيانات التي تصفها معينات البحث.

(١) يسمى هذا العنصر " وصف المكونات التابعة " [المترجم].

وفـي هـذه الحالـة الأخيـرة يـتم اسـتخدام العنـصر الخـاص المـسمى " كـيان أرشـيفي رقمـي <dao> . وإذا وجدت عدة تمثيلات رقمية للكيان نفسه (مثل النسخ المنمنمة أو المصغرة thumbnail والصورة في صيغة جي بيج JPEG مثلاً) يتم جمع الروابط للإصدارات المختلفة داخل العنصر الغلاف المسمى <daogroup> [١]. هكذا:

```
<c٠٣ level="item"><did>
        <unittitle>Letter to Doreothea Huxley</unittitle>
        <unitdate>May ٤, ١٩٢٩</unitdate>
        <dao href="http://www.server.edu/letter١٢٤.jpg"></dao></did>
</c٠٣>
```

ومـن المفتـرض استخدام الوصـف الأرشـيفي المرمـز مـع نمـاذج إخـراج الـصفحات Stylesheet لإنـشاء عـروض مطـبوعة أو مـتاحة علـى الخـط المباشر لمعينات البحث ، مع تضمين معرف نوع الوثيقة DTD على العديد مـن العناصـر التـي تيسر عملية العرض ، ويتيح عنصر الهوية أو التعريف الوصـفي <did> ومعظـم العناصـر الأخـرى داخـل قسم الوصـف الأرشـيفي <archdesc> بتـضمين العنـصر الفرعـي المسمى " الترويسة " <head> مـن أجـل إدخـال ترويسة القسم ، كمـا تتيح معظم العناصر أيضاً إمكانية إضافة الخاصـية المسماة " الوسـيمة" <label> ، بحيث يمكن للترميز المسجل على النحو التالي :

```
<did>
```

(١) يسمى هذا العنصر " مجموعة الكيان الأرشيفي الرقمي " [المترجم].

<repository label="Repository">The Chester A. Mann Archives</repository>

.....

أن ينتج عروضاً مختلفة بناء على نموذج الإخراج المستخدم ، هكذا:

Repository: The Chester A. Mann Archives

أو :

Repository. The Chester A. Mann Archives

ونظراً لأن معينات البحث المطبوعة غالباً ما تعرض قوائم الحاويات الخاصة بالصناديق والمجلدات وبكرات المواد المصغرة وغيرها في شكل جدولي، هناك أيضاً عناصر خاصة لتيسير هذا العرض الجدولي.

وبحكم أن الوصف الأرشيفي المرمز كان قد صمم لدعم عمليات الربط mapping بين ما وراء البيانات في الوصف الأرشيفي المرمز وصيغة مارك ، فإن معظم العناصر تتيح إمكانية إضافة الخاصية المسماة " المناظر الترميزي encodinganalog" التي تأخذ قيمة تاج مارك على النحو التالي :

<subject source="lcsh" encodinganalog="٦٥٠">Civil War-Forida</subject>

ونظرياً يمكن تحويل الوصف الأرشيفي المرمز عن طريق أحد البرامج الآلية إلى تسجيلة مارك باستخدام خصائص " المناظر الترميزي encodinganalog"، مع أن تسميات المحتوى designating content داخل حقول مارك (أي قيم المؤشرات values indicators وتقنينات الحقول الفرعية subfieldings) ستظل محاطة بالإشكاليات. وبالعكس فإنه يمكن

استخدام تسجيلات مارك الحالية الخاصة بإحدى المجموعات لتسكين عناصر معينة داخل الوصف الأرشيفي المرمز. وبالرغم من أن مثل هذه التحويلات نادراً ما تنفذ عملياً ، إلا أنه من الشائع الربط بين تسجيلة فهرسة على مستوى المجموعة في صيغة مارك والوصف الأرشيفي المرمز للمجموعة نفسها، كما يمكن أن يتم تحويل الوصف الأرشيفي المرمز إلى صيغة لغة ترميز النص الفائق HTML وتخزينه بها ، كما يتم تسجيل المعرف الموحد للمصدر URI الخاص بالمصادر للوصف الأرشيفي المرمز في الحقل "٨٥٦" في صيغة مارك. هكذا:

٨٥٦ ٤٢ $٣ Finding aid for this collection

$u http://www.server.edu/archives/fa١٢٣٤.html

وبالرغم من أن الوصف الأرشيفي المرمز يعنى بتعريف الجوانب الدلالية لمعينة البحث، إلا أنه لم يكن هناك حتى وقت قريب معايير محتوى متوافقة مع معينات البحث ، ويعد الإطار الأكثر عمومية للوصف الأرشيفي هو ما يعرف "بالتقنين الدولي العام للوصف الأرشيفي the General International Standard Archival Description (ISAD(G)) " الذي طورته لجنة المعايير الوصفية التابعة للمجلس العالمي للأرشيفات the International Council on Archives Committee on Descriptive Standards [١]، حيث

(١) International Council on Archives, "ISAD(G): General International Standard Archival Description: adopted by the Committee on Descriptive Standards, Stockholm, Sweden, ١٩-٢٢ September ١٩٩٩," ٢nd ed., available at

http://www.ica.org/biblio/com/cds/isad_g_٢e.pdf. Accessed ٢٤ June ٢٠٠٢.

يوفر هذا المعيار (ISAD(G)) وصفاً عالي المستوى لستة وعشرين عنصراً يمكن استخدامها في تسجيلات الفهرسة على مستوى المجموعة ومعينات البحث ، غير أنه يعتمد على معايير أخرى لتوفير قواعد المحتوى. وفي كندا تم تنفيذ قواعد الوصف الأرشيفي Rules for Archival Description (RAD) التي نشرت أول مرة عام ١٩٩٠م، ويتم استخدامها على نطاق واسع كدليل إرشادي لكيفية صياغة محتوى معينات البحث[١]. وفي الولايات المتحدة بالتحديد قدم تطوير الوصف الأرشيفي المرمز حافزاً لمجتمع الأرشيفيين لدراسة ممارسات إنشاء معينات البحث والعمل على التطوير المستمر لمعايير المحتوى الخاصة بها، كما تعمل جمعية الأرشيفيين الأمريكيين مع المجلس الكندي للأرشيفات على مشروع التقارب بين المعايير الأمريكية والكندية المعروف باسم "كسترد" CUSTARD. وحاول هذا المشروع تطوير إصدارة ذات صيغة عالمية من قواعد الوصف الأرشيفيRAD لتحقيق التناغم بين الممارسة الأمريكية والكندية فضلاً عن توحيد إرشادات المحتوى الخاصة بكل من فهرسة معينات البحث وإنشائها.

وتعد أكبر عمليات الوصف الأرشيفي المرمز تلك التي تنفذ في أقسام المجموعة الخاصة في المكتبات الأكاديمية، حيث تتواجد الآن مجموعات للوصف الأرشيفي المرمز في كل من جامعة هارفارد (http://findingaids.harvard.edu/) ، ومكتبة جامعة فرجينيا (http://www.lib.virginia.edu/speccol/guides/) ، وجامعة ديوك(

---

(١) Canadian Committee on Archival Description, Rules for Archival Description, available at http://www.cdncoucilarchives.ca/archdesrules.html. Accessed ٢٦ August ٢٠٠٢.

http://odyssey.lib.duke.edu/findaid/) والعديد من الجامعات الأمريكية الأخرى. كما أنشأت بعض الولايات ، مثل "كينتاكي" و"نيومكسيكو" قواعد بيانات موحدة لمعينات البحث ساهمت في تكوينها مؤسسات أكاديمية وجمعيات تاريخية ومكتبات حكومية. كما تقوم إحدى أكبر المجموعات المتمثلة في أرشيف كاليفورنيا المباشر Online Archive of California بتجميع معينات البحث المرمز وفقاً للوصف الأرشيفي المرمز EAD-encoded من الأرشيفات والمتاحف والمكتبات من جميع أنحاء ولاية كاليفورنيا(http://oac.cdlib.org/).

وبالرغم من ذلك ، فإن الهيئات الحكومية الفيدرالية والجمعيات التاريخية وبعض المؤسسات الأخرى كانت بشكل عام أبطأ في تبني الوصف الأرشيفي المرمز من الهيئات الأرشيفية التابعة لمؤسسات أكاديمية ، وكانت حجة البعض في ذلك أن الوصف الأرشيفي المرمز لا يناسب وصف تلك الأنواع الأخرى من المجموعات ، مع أن البعض الآخر لم يجد فارقاً جوهرياً في الوصف الأرشيفي [حتى مع تفاوت] طبيعة المؤسسة الأم. ولا شك أن الوصف الأرشيفي المرمز يعد واحداً من أصعب معايير ما وراء البيانات في التنفيذ لأنه يتطلب خبرة في الوصف الأرشيفي مع معرفة معقولة بمعرف نوع الوثيقة في صيغتي لغة الترميز المعيارية العامة ولغة الترميز الموسعة SGML/XML DTD المعقد ، كما أن تنفيذه يتطلب أدوات تحرير وبرمجيات بحث وعرض النصوص المرمزة بصيغتي لغة الترميز المعيارية العامة ولغة الترميز الموسعة SGML/XML-encoded texts التي قد تجد العديد من المؤسسات الصغيرة صعوبة في دعمها. وقد كان هذا عبئاً ثقيلاً بشكل خاص عندما كان الوصف الأرشيفي المرمز قائماً على صيغة لغة الترميز المعيارية العامة SGML-based ؛ حيث لم يكن هناك

سوى القليل من التطبيقات البرمجية المصممة لدعم النصوص المرمز بهذه اللغة. ويبدو أنه مع الشهرة التي تحظى بها لغة الترميز الموسعة XML سيكون هناك مستوى أعلى من الألفة للترميز اعتماداً على هذه اللغة ، وستكون هنالك فرص أوسع لاختيار الأدوات التي تدعم إنشاء واسترجاع وعرض ما وراء البيانات المعتمدة على لغة الترميز الموسعة XML-based ، حتى باتت الاتجاهات التي تدعم الوصف الأرشيفي المرمز راسخة في عدد كبير من الأرشيفات.

READINGS :

The EAD is extremely well documented. The SAA publishes the tag library and a very useful manual of application guidelines, both of which are available in print and on the Web:

- Encoded Archival Description (EAD) Official Web Site. Accessed ٢٦ August ٢٠٠٢. Available at http://www.loc.gov/ead.

  The Library of Congress is the official maintenance agency for EAD documentation. Their EAD website contains background information and pointers to official versions of the DTDs and documentation.

- Society of American Archivists, EAD Round Table. "EAD Help Pages." Accessed ٢٤ June ٢٠٠٢. available at http://www.iath.virginia.edu/ead/.

  A compilation of links to additional EAD documentation and tools. A particularly useful set of tools, including data entry templates for various SGML/XML authoring software and stylesheets for displaying EAD, is available in the EAD Cookbook.

- Society of American Archivists. Encoded Archival Description: Application Guidelines. Version ١,٠. Chicago: SAA, ١٩٩٩. Available at http://Icweb.loc.gov/ead/ag/aghome.html . Accessed ٢٤ June ٢٠٠٢.

- Society of American Archivists. Encoded Archival Description: Tag library. Version ١,٠. Chicago:SAA, ١٩٩٨. Available at http://Icweb.loc.gov/ead/tglib/tlhome.html . Accessed ٢٤ June ٢٠٠٢.

  A wealth of articles has been published about all aspects of EAD development and implementation. American Archivist ٦٠ (fall

١٩٩٧) is a special issue devoted to the EAD. Articles of particular interest from this and other sources include:

- Fox, Michael. "Implementing Encoded Archival Description: An Overview of Administrative and Technical Considerations."

The practical details of software and systems to support EAD.

- Kiesling, Kris. "EAD as an Archival Descriptive Standard."

Describes the EAD in context of other standards efforts.

- Meissner, Dennis. "First Things First: Reengineering Finding Aids for Implementation of EAD."

How the EAD has caused archivists to reevaluate the form and function of finding aids.

- Pitti, Daniel. "Access to Digital Representations of Archival Materials: The Berkeley Finding Aid Project." RLG Digital Image Access Project: Proceedings from an RLG Symposium (Palo Alto: The Research Libraries Group, ١٩٩٥),٧٣-٨١. Available at http://sunsite.berkeley.edu/FindingAids/EAD/diap.html.

The early development of the EAD.

- Smith, MacKenzie. "DFAS: The Distributed Finding Aid Search System." Lib Magazine ٦, no. ١ (January ٢٠٠٠).

Available at http://www.dlib.org/dIib/januaryOO/Olsmith.html.

A project to implement broadcast search across distributed repositories of EADs

الفصل العاشر

ما وراء البيانات لمجالي الفنون والعمارة

حظـي وصـف الأعمـال الفنية والمعمارية وغيرها من الأعمال البصرية باهـتمام معهـد جيتـي للمعلـومات Getty Information Institute، وجمعـية المصادر البصرية the Visual Resources Association (VRA) ، فضلاً عن كثير من الأفراد والهيئات الضالعين في مجال الدراسات العلمية للفنون . ونلقي الضوء في هذا الفصل على ثلاثة جهود في هذا الصدد، هي: الفهرسة التقليدية اعتمادًا على صيغة مارك، وفئات وصف الأعمال الفنية، والفئات الأساسية لجمعية المصادر البصرية .

فهرسة المواد البصرية :

يمكـن فهرسـة المـواد البـصرية (أي الرسـوم الفنـية، الـصور الفوتوغرافية، الرسوم التصويرية graphic images) اعتمادًا على قواعد الفهرسة التقليدية المتبعة في المكتبات، وقد طورت صيغة مارك للمواد البصرية في ثمانينات القرن العشرين الميلادي، وحتى بعد الاستغناء عن الصيغ المستقلة لمارك والاعتماد على الصيغة التكاملية في التسعينات، ظلت عناصر البيانات الخاصة بالمواد البصرية كما هي ولم يستغن عنها. وقد خصص الفصل الثامن من القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية للرسوم التصويرية - التي تشمل المواد غير الشفافة opaque object - كالأعمال الفنية الأصلية ثنائية الأبعاد ومستنسخاتها، والمواد المخصصة لأغراض العروض كالشرائح، ومجموعات الرسوم التصويرية . كما خصص الفصل العاشر لفهرسة أعمال النحت والمجسمات الطبيعية والصناعية وغيرها من الأعمال الفنية ثلاثية الأبعاد Three- dimensional Artifacts and Realia ، كما تعالج القواعد التي يتضمنها الفصل الرابع المواد

غير المنشورة، والمجموعات الأرشيفية .

كذلك أصدرت مكتبة الكونجرس دليلاً إرشاديًا بعنوان " Graphic Materials- Rules for Describing Original Items and Historical Collection " لفهرسة المواد التصويرية ، في إطار القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية، وإن كان يتضمن قدرًا من المرونة التي تتيح إمكانية الإضافة والتعديل عما تتضمنه القواعد في الحالات التي تستلزم ذلك(١).

وإذا كان النجاح قد حالف كثيرًا من المشروعات التي طبقت القواعد الأنجلو أمريكية للفهرسة وصيغة مارك في فهرسة المواد البصرية، فلا يعني ذلك أن خطى التطبيق، وخاصة ما يتصل منها بمارك وبقواعد المحتوى، لم تعرقلها تحديات كثيرة ؛ فالأعمال الفنية الأصلية غالبًا ما تحفظ في المتاحف وصالات العرض التي تفتقر إلى المتخصصين الببليوجرافيين الذين تتوافر لديهم الخبرة الكافية لإنشاء تسجيلات مارك بكفاءة، كما أن مستنسخات الأعمال الفنية (مثل شرائح العرض والصور) قد تقتنى في بعض الأقسام الأكاديمية أو المكتبات المتخصصة دون خبرة بفهرستها فهرسة تقليدية، ومع أن استخدام القواعد الأنجلو أمريكية للفهرسة يفترض مسبقًا توافر مقومات وصف المواد النصية والبصرية على حد سواء، فهنالك بلا شك اختلافات جوهرية بين وصف الكتب التي عادة ما تنشر في أعداد متشابهة من النسخ التي لا تكاد تخلو أي منها من صفحة عنوان رئيسية،

(١) Elisabeth Betz Parker, Graphic Materials-Rules for Describing Original Items and Historical Collection, ١٩٨٢ with ١٩٩٦ updates, available at

http://www.tIcdelivers.com/tic/crs/grph٠١٩٩.html. Accessed ٢٥ June ٢٠٠٥. See.

والأعمال الفنية التي قد يكون من النادر توافر أصولها بين يدي القائم على فهرستها، أو لا تتوافر له المعلومات الموثوقة عنها .

ولا شك أن تطبيق القواعد الأنجلو أمريكية للفهرسة تحيط به بعض الإشكالات من عدة جوانب، منها أن كلا الأصول و المستنسخات أو بدائلهما تتطلب أن توصف مستقلة بعضها عن بعض، ومن الإشكالات كذلك أن كثيرًا من المواد التي توصف تفتقر إلى مصادر المعلومات التي يعتمد عليها في استقاء معلومات الوصف، ولعل من الأمثلة على ذلك أن الفصل الثامن من الطبعة المنقحة من القواعد الأنجلو أمريكية للفهرسة (AACR٢R) في طبعتها الثانية لم يعط أمثلة كافية لكيفية صياغة عناوين للمواد التي تفتقر إلى عناوين ، واكتفى بمثالين فقط على هذه الحالة ، جاء أحدهما هكذا : [photograph of Alice Liddell صورة فوتوغرافية لـِ " أليس لايدل"][(١)] ، وواضح أن ما لهذا المثال سوى إفادة قليلة لمن يحاول فهرسة مجموعة من الصور الفوتوغرافية ، ما لم يكن من المفضل أن ترتب جميع العناوين تحت المدخل " Photograph of... ، صورة لـِ ... " .

إن قضية كيفية وصف رسمات الأعمال الفنية أو المعمارية لهي قضية معقدة، بل إن مسألة تسمية تلك الرسوم يكتنفها قدر من الغموض ؛ فمصطلح "بديل surrogate" يوحي بأن الكيان الثاني قصد به أن يكون بديلاً عن الكيان

(١) يعرف هذا في سياق الفهرسة بالعنوان المصنوع ، حيث يلجأ المفهرس في حالة افتقار العمل لصفحة العنوان أو عدم وجود عنوان إلى عمل أو اقتراح عنوان للمادة ، ويحدث ذلك كثيراً في حالة الصور الفوتوغرافية للأشخاص أو الأماكن ، وفي هذه الحالة يبدو أن العمل عبارة عن صورة لهذا الشخص ومن ثم صيغ العنوان المصنوع هكذا [المترجم] .

الأصلي ، وذلك ليس صحيحًا في جميع الأحوال، كما أن مصطلح " مستنسخ reproduction " يوحي الالتزام الآلي بالأصل، وهو ما لا ينطبق على الكثير من الصور الفوتوغرافية ، ذلك أن استنساخ (الأعمال) الفنية art reproduction عرف في ملحق القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية في كلمات محددة تقول :" هو نسخة من العمل الفني أنتجت إنتاجا آليًا ، وينظر إليها عامة وكأنها أشبه بالطبعة التجارية " . وقد كانت جمعية المصادر البصرية تفضل في البداية استخدام مصطلح " وثيقة بصرية visual document " لكنها رأت بعد ذلك أن تستخدم مصطلح " صورة بصرية visual image " ، أو مصطلح " صورة image " فقط ، ولأغراض التبسيط رأينا في هذا السياق أن نستخدم مصطلح "تمثيل بصري visual representation " أو مصطلح " تمثيل representation " فقط.

وبغض النظر عن المسميات التي تتخذها تلك التمثيلات ، فإن أسلوب التعامل معها هو الذي يحظى بالاهتمام عند أية محاولة لوصف مجموعة الأعمال الفنية والمعمارية ؛ فالمتاحف وصالات العرض التي تمتلك أعمالاً أصلية عادة ما تحتفظ بمجموعة من التمثيلات لهذه الأعمال ، والتي قد تكون مجرد صور فوتوغرافية لها ، أو حتى صورًا مأخوذة بالأشعة السينية ، كما أن مجموعات المصادر البصرية المستخدمة في الأغراض التعليمية ، قد لا يكون من بينها أعمال أصلية على الإطلاق، بل مجرد مجموعة من الصور الفوتوغرافية، أو الشرائح، أو الصور الرقمية، أو كل هذه الأشكال، وبصرف النظر عن الذي تسعى إليه قواعد الفهرسة، تحاول تسجيلات مارك للتمثيلات البصرية الجمع بين المعلومات الخاصة بالعمل الفني نفسه،

والمعلومات الخاصة بالتمثيلات الخاصة به، ولكن بطرق غير مقننة .

والواقع أنـه لـيس هناك توحيد لـرؤى المفهرسين في استخدام مارك في فهرسة هذه المواد، إما بسبب صعوبة تطبيق قواعد الفهرسة، أو لعدم توافر الأمثلة والإرشادات الكافية، أو لغير ذلك من أسباب أخرى . وقد أكد ذلك " الدليل الإرشادي لتطبيق مارك الفنـي ArtMarc Sourcebook " الـذي قارن بين تطبيق مارك في ثلاثة وعشرين مشروعًا لوصف أعمال فنية أصلية أو تمثيلات لها أو لكليهما ، وتبين لـه أن ثمة تفاوتًا كبيرًا بين تلك التطبيقات . ومع عدم الاطراد أو التجانس فـي استخدام جميع حقول مارك تقريبًا، فقد أحاطت بتطبيق بعض الحقول كثير من الإشكالات ، منها على سبيل المثال، التساؤل عن ماهية الحقل الـذي ينبغـي أن تسجل فيه المعلومات المتصلة بالتمثيل البصري ، وعن أي من حقول التبصرات ما يمكن استخدامه للمعلومات غير المعرفة بشكل محدد في مارك ، وعن الكيفية التي تستخدم بها الحقول الموضوعية . وفضلاً عن تلك التفاوتات فقد بدا فـي كثير من الحـالات أن ثمـة فـوارق ثقافـية بـين إحـصائيي المكتـبات مـن جهـة، واختصاصيي المصادر البصرية من جهـة أخرى ، فـي منظورهم لهذه المواد، [وقد أشار إلى ذلك بعض الباحثين بقوله] :

" إن المكتبيين الذين يفهرسون الرسوم اليدوية والصور الفوتوغرافية للمباني يعتبرون أن أسماء المباني بمثابة موضوعات أو أسماء، أو هيئات اعتبارية ؛ ذلك أنهم يفهرسون تمثيلات تلك المباني وليس المباني ذاتها، أما مفهرسو المصادر البصرية - في المقابل - فيفهرسون العمل بشكل عام،

وهو في هذه الحالة المبنى نفسه، وليست تلك الشريحة الفيلمية لهذا المبنى أو صورة أو رسمًا معماريًا له، ومن ثم يكون العنوان هو اسم ذلك المبنى(١).

وتعتري فهرسة المصادر البصرية مشكلة أخرى تتمثل في مدى استخدام اللغات المقيدة التي تناسب وصف الأعمال الفنية والمعمارية ، فهنالك شعور بأن قوائم رؤوس الموضوعات العامة - مثل قائمة رؤوس موضوعات مكتبة الكونجرس - لا تصلح للتطبيق في هذا المجال ، ويؤكد ذلك اعتماد العديد من مجموعات المصادر البصرية في وصفها على قوائم داخلية وتجاهلهم لأية قوائم استنادية أخرى . وقد برزت منذ ثمانينات القرن العشرين الميلادي اتجاهات عملية في هذا الصدد ، لعل أبرزها ما قام به معهد جيتي للمعلومات ، إذ تكفل بتطوير ثلاثة أعمال أصبحت جميعها من أدوات العمل الأساسية لوصف المصادر البصرية ، وهي : مكنز الفنون والعمارة the Art and Architecture Thesaurus (AAT) ، والقائمة الموحدة لأسماء الفنانين the Union List of Artist Names (ULAN)، ومكنز جيتي للأسماء الجغرافية Getty Thesaurus of Geographic Names (TGN). كذلك طورت بعض الجهات الأخرى أعمالاً مماثلة ، منها " مكنز المواد التصويرية Thesaurus for Graphic Materials ١: Subject Terms " - الذي أصدرته مكتبة الكونجرس خصيصًا للاعتماد عليه في التكشيف الموضوعي "للصور التاريخية المتوافرة في العديد من المكتبات والجمعيات التاريخية ، والأرشيفات والمتاحف" ، ومن هذه الأعمال كذلك تصنيف أيكون ICON-

(١) Linda McRae and Lynda S. White, eds., ArtMARC Sourcebook (Chicago: American Library Association, ١٩٩٨), ١٠.

CLASS ، أحد النظم العالمية للتصنيف الموضوعي للصور الفنية ، وأخيراً تجدر الإشارة في هذا الصدد إلى أن استخدام أدوات ضبط المنفردات كهذه يحظى بتشجيع خطط ما وراء البيانات المعنية بوصف الأعمال الفنية والمعمارية .

فئات وصف الأعمال الفنية **(CDWA)** :

طورت لجنة المهام لمعلومات الفنون Art Information Task Force (AITF) في بداية التسعينات من القرن العشرين " فئات وصف الأعمال الفنية "Categories for the Description of Works of art (CDWA)" ، وهو مشروع تبنته كل من جمعية مكتبات الفنون الأمريكية ، وبرنامج جيتي لمعلومات تاريخ الفن - الذي عرف فيما بعد باسم معهد جيتي للمعلومات - وكانت هذه اللجنة قد نظمت لقاءً جمع نخبة من مؤرخي الفنون ، وأمناء المتاحف ، ومسؤولي تسجيل المواد المتحفية، واختصاصيي المصادر البصرية، بهدف التوصل إلى اتفاق وجهات نظر جميع هؤلاء المهتمين حول عناصر أساسية لوصف الأعمال الفنية . وقد تركز اهتمام هذا العمل في أول الأمر على الكيانات " المتحركة أو القابلة للنقل movable " التي قامت المتاحف على جمعها، وعلى تمثيلات هذه الكيانات، وعلى أعمال الفن الاستعراضي (ثم أضيف إليها بعد ذلك أعمال الفن المعماري). وفي أثناء انخراط تلك اللجنة في هذا الأمر ، قدمت نظم إدارة المجموعات المتحفية أداة مبدئية لضبط هذه المجموعات inventory control ، غير أن الباحثين لم يروا استخدامها في أغراضهم البحثية إلا بشكل محدود ، ومن ثم تركز اهتمام اللجنة بوضوح على تشجيع الاتجاه نحو تسجيل المعلومات بأسلوب يكفل الإفادة منها في سياق البحث العلمي المعني بتاريخ الفنون . وقد

صدرت الإصدارة الأولى من [خطة] " فئات وصف الأعمال الفنية CDWA " في عام ١٩٩٤م، ثم تبعتها الإصدارة التالية في عام ٢٠٠٠م، وهي الإصدارة الحالية "رقم ٢,٠ " ، التي تتاح على موقع معهد جيتي للمعلومات[١].

وتعرف هذه الخطة الفئات الدلالية مع بعض قواعد المحتوى، دون تحديد أي تركيبة معينة لهذه الفئات ، وقد كان هناك افتراض بأن تعرض جميع البيانات في جداول قواعد بيانات علاقية ، ولكن في الوقت نفسه ليس من الملزم أن تقوم قواعد البيانات المحلية بتطبيق الفئات كما تحدد في سياق الخطة بالضبط . وقد ورد بمقدمة الخطة ما يفيد بذلك ، حيث جاءت الفقرة التالية:

" تصنف الفئات[٢] محتوى قواعد بيانات الفنون من خلال وضع إطار فكري لوصف الكيانات والصور وإتاحة المعلومات عنها ، كما تسعى هذه الفئات إلى التعريف بكل من مصادر المفردات vocabularies resources والممارسات الوصفية بالشكل الذي يؤدي إلى أن تصبح المعلومات المتواجدة في النظم المختلفة[٣] أكثر توافقًا وأكثر قابلية للوصول إليها ، كما أنها تقدم كذلك إطارًا يمكن أن تربط به نظم المعلومات الفنية القائمة ، كما يمكن اعتماداً عليه تطوير نظم جديدة" .

---

(١) Categories for the Description of Works of Art, edited by Murtha Baca and Patricia Harpring. available at http://www.getty.edu/research/institute/standards/cdwa. Accessed ٢٥ June ٢٠٠٤.

(١) أي فئات وصف المصادر البصرية[المترجم] .

(٢) يقصد بها نظم المعلومات المتخصصة في مجال الفنون [المترجم].

وتحدد خطة " فئات وصف الأعمال الفنية " سبعًا وعشرين فئة أساسية ، تتبع كل منها فئاتٌ فرعية أخرى متعددة بحيث تنتج في جملتها ما يقارب ثلاثمائة عنصرٍ ، حُدد نحو أربعة وعشرين عنصرًا منها كعناصر " أساسية core " اعتبرت هي الحد الأدنى لمجموعة العناصر اللازمة لتعريف العمل الفني بشكل فريد لا لبس فيه . ولكن ينبغي ملاحظة أن تغيرًا جوهريًا قد حدث بين الإصدارتين اللتين صدرتا من هذه الخطة ؛ حيث قسمت الفئات في الإصدارة الأخيرة رقم ٢,٠ " إلى مجموعتين : تعرف أولاهما بمجموعة " الكيان ، أو البنية المعمارية أو المجموعة Object, Architecture or Group " ، وتعرف ثانيتهما بمجموعة " القوائم الاستنادية أو ضبط المفردات Authorities /Vocabulary control " . على أن تخصص المجموعة الأولى لوصف العمل نفسه (أو تمثيله ، أو تسجيلة ما وراء البيانات التي تصفه)، وتخصص المجموعة الثانية " لوصف المعلومات الشكلية عن الأفراد ، والأماكن والمفاهيم المتصلة بالعمل ، بدعوى أن هذه البيانات يفضل اختزانها في تسجيلات استنادية .

وقد حددت خمس خصائص لتعريف الفئات، جاء ت على النحو التالي : خاصية " التعريف Definition " ، وخاصية " المناقشة Discussion "، وخاصية "العلاقات Relationships " ، وخاصية "الاستخدامات Uses "، وخاصية "الإتاحة أو الوصول Access" مع إمكانية إضافة الخاصيتين : " الأمثلة Examples "، "والمصطلح /الصيغة Terminology/format " للفئات الفرعية . ويعرض الشكل رقم (١/١٠) تعريفًا لإحدى الفئات الفرعية المسماة " المقاييس - الشكل الخارجي Measurements- Shape " .

ومن الملاحظ أن اختيار هذه الخصائص وتعريفها يؤكد تركيز مؤرخي الفنون والباحثين في هذا المجال على جانب استخدام هذه المصادر؛ فخاصية " المناقشة " هي عبارة عن تفسير سردي للكيفية التي تستخدم بها الفئة ، ويشمل ذلك الأهمية التاريخية الفنية للبيانات، أما خاصية " الاستخدامات " فتوضح الكيفية التي يمكن للباحث أن يستخدم بها البيانات أو يفيد منها، حيث يشار- على سبيل المثال - بالنسبة للعنصر المسمى " العنوان أو الأسماء - نص Title or Name- Text " إلى ما يلي : " في بعض الحالات يقدم العنوان الذي اختاره الفنان للعمل إيضاحات مهمة عن فحوى العمل "، أما خاصية " الإتاحة أو الوصول " فتبين الكيفية التي تستخدم بها الفئة في سياق عملية الاسترجاع ، وتبين تحديدًا ماهية البيانات التي تستخدم كنقطة إتاحة أساسية .

ولعل الاعتراف الصريح بعدم يقينية uncertainty ، وعدم موضوعية المعلومات التاريخية الفنية بشكل عام هو من أبرز ملامح " فئات وصف الأعمال الفنية " ، يضاف إلى ذلك أن القدرة على تسجيل الأسماء المتغيرة للأفراد والأماكن والموضوعات تقع هي الأخرى في بؤرة اهتمام هذه الفئات ، ناهيك عن

Measurements – Shape

DEFINITION

[illegible]

DISCUSSION

[illegible]

TYPES

[illegible]

ACCESS

[illegible]

TERMINOLOGY/FORMAT

[illegible]

الشكل رقم (١٠/١)

تعريف الفئة الفرعية " المقاييس – الشكل الخارجي " من الفئات الأساسية لوصف المصادر البصرية

المصدر :

Baca, Murtha, and Patricia Harpring, eds. Categories for the Description of Works of Art. The J. Paul Getty Trust and College Art Association, ٢٠٠٠. Available at http://wiww.getty.edu/research/institute/standards/cdwa/.

أن جميع الفئات تقريبًا تتيح إمكانية إضافة الفئتين الفرعيتين "الملحوظات الجوهرية Remarks " - التي يمكن استخدامها لتدوين الملحوظات البحثية التي تشبه الحواشي التي ترافق الأعمال البحثية، و"الإرجاعات الببليوجرافية Citations " - التي تستخدم لتوثيق المصدر الذي أخذت منه جميع المعلومات في [تسجيلة] ما وراء البيانات . ومن الواضح أن مثل هذه الفئات الفرعية تعين الباحث على تقدير مدى دقة المعلومات المتاحة، ومدى مصداقيتها .

وتشتمل النسخة المتاحة على الإنترنت من خطة "فئات وصف الأعمال الفنية" على أمثلة كثيرة من الوصف المصدري لعدة أنواع مختلفة من الكيانات ، مثل : الكيانات المطبوعة والمرسومة، والصور الفوتوغرافية، وأعمال النحت وأعمال التطريز، والمطبوعات اليدوية، والمجسمات . ويعرض الشكل رقم (٢/١٠) مثالاً لفهرسة أحد الرسوم الزيتية .

ولتعقد هذا العمل واتساعه ، فمن النادر تطبيقه تطبيقًا كاملا ، ومع ذلك فإن التعامل معه كإطار عمل أهله لأن يستخدم أساساً لعدد كبير من قواعد البيانات المتحفية ، كما أنه أدَّى دورًا مؤثرًا في تطوير مواصفات ما وراء البيانات الخاصة بكثير من المشروعات والتطبيقات ، التي من أهمها : " نقاط الإتاحة لمشروع المعلومات المباشرة للتراث الحضاري[١] CIMI's CHIO " التي تبناها ائتلاف التبادل الآلي للمعلومات المتحفية في منتصف عقد التسعينات من القرن

CDWA Fielded Example: Oil Painting 1

Example with Image
Next Fielded Example >>

1.1	Object/Work - Type	painting
2.1	Classification - Term	Paintings
4.1	Title or Name - Text	Irises
7.1	Measurements - Dimensions	71 x 93 cm (28 x 36 5/8 in.)
7.1.2	Measurements - Dimensions - Type	height
7.1.3	Measurements - Dimensions - Value	71
7.1.4	Measurements - Dimensions - Unit	cm
7.1.2	Measurements - Dimensions - Type	width
7.1.3	Measurements - Dimensions - Value	93
7.1.4	Measurements - Dimensions - Unit	cm
8.1	Materials and Techniques - Description	oil on canvas, applied with brush and palette knife
8.3.1	Materials and Techniques - Materials - Name	[illegible]
8.3.2	Materials and Techniques - Processes - Implement	brush
8.3.2	Materials and Techniques - Processes - Implement	palette knife
8.4.1	Materials and Techniques - Materials - Name	oil paint
8.4.1	Materials and Techniques - Materials - Name	canvas
14.1	Creation - Creator	Vincent van Gogh
14.1.3	Creation - Creator - Identity	Gogh, Vincent van (Dutch painter, 1853-1890)
14.1.4	Creation - Creator - Role	painter
14.2	Creation - Date	1889
14.2.1	Creation - Date - Earliest Date	1889
14.2.2	Creation - Date - Latest Date	1889
18.1.1	Subject Matter - Description - Indexing Terms	irises
18.1.1	Subject Matter - Description - Indexing Terms	landscape
18.1.1	Subject Matter - Description - Indexing Terms	soil
18.1.1	Subject Matter - Description - Indexing Terms	nature
18.2.1	Subject Matter - Identification - Indexing Terms	Irises
18.3.1	Subject Matter - Interpretation - Indexing Terms	regeneration
19.1	Context - Historical/Cultural	exhibited at Salon des Indépendants, September 1889
19.1.2	Context - Historical/Cultural - Event Name	Salon des Indépendants
19.1.3	Context - Historical/Cultural - Date	September 1889
26.1	Current Location - Repository Name	J. Paul Getty Museum
26.2	Current Location - Geographic Location	Los Angeles (California, USA)
26.3	Current Location - Repository Numbers	90.PA.20
27.1	Descriptive Note - Text	This work was painted when the artist was recuperating from a severe attack of mental illness, and it depicts the garden at the asylum at Saint-Rémy. It is influenced by the work of Gauguin and Bernard, and is remarkable for the contrasts of color.
28.1	Creator Identification - Name	Gogh, Vincent van
28.1	Creator Identification - Variant Names	Gogh, Vincent Willem van
28.1	Creator Identification - Variant Names	van Gogh, Vincent
28.1	Creator Identification - Variant Names	Vincent van Gogh
28.3	Creator Identification - Dates/Locations	1853-1890, active in Holland
28.3.1	Creator Identification - Birth Date	1853
28.3.2	Creator Identification - Death Date	1890
28.3.3	Creator Identification - Places of Activity	Holland
28.3.3	Creator Identification - Places of Activity	Netherlands, the
28.4.1	Creator Identification - Nationality/Citizenship	Dutch
28.6	Creator Identification - Life Roles	painter

(١) ...on
ائتلا...
أساسيات...

نموذج فهرسة مأخوذ من الإصدارة المتاحة على الوب من فئات وصف الأعمال الفنية

المصدر :

Baca, Murtha, and Patricia Harpring, eds. Categories for the Description of Works of Art. The J. Paul Getty Trust and College Art Association, ٢٠٠٠. Available at http://wiww.getty.edu/research/institute/standards/cdwa/

العشرين، ودليل شبكة تداول المواد المتحفية the Museum Loan Network Directory. كذلك تحتفظ النسخة المتاحة على الوب من هذه الخطة بقائمة بالمعايير المتصلة بها ، أو التي تشكل أساسًا لها، من أهمها : دليل المجموعة الاستشارية لمؤسسة الوثائق والرسومات المعمارية الموسوم " الدليل الإرشادي لوصف فئات بيانات الرسوم المعمارية Guide for the description of Architectural Drawings data categories " ، و"معجم ائتلاف صور متاحف الفنون the Art museum Image Consortium(AMICO) "، " ونقاط الوصول الخاصة بمبادرة ائتلاف التبادل الآلي للمعلومات المتحفية "، "الدليل الإرشادي الدولي لمعلومات الكيانات المتحفية للجنة الدولية للتوثيق التابعة للمجلس الدولي للمتاحف CIDOC's International Guide for Museum Object Information"، و"توليفة أو طيف جمعية التوثيق المتحفي the Museum Documentation Association (MDA)Spectrum " ، و" الفئات الأساسية لجمعية المصادر البصرية the Visual Resources Association (VRA)".

الفئات الأساسية لجمعية المصادر البصرية **VRA** :

اضطلعت لجنة معايير البيانات التابعة لجمعية المصادر البصرية بتطوير مواصفة الفئات الأساسية لجمعية المصادر البصرية ، وعندما بدأت اللجنة عملية التطوير هذه ، كانت " فئات وصف الأعمال الفنية " متداولة

بين المهتمين لمراجعتها ، وكان أعضاء هذه اللجنة على دراية بطبيعة هذا العمل ، ومع ذلك لم يكن مفهرسو المصادر البصرية على القدر نفسه من الاهتمام الذي كان عليه مؤرخو الفنون وأمناء المتاحف الذين طوروا فئات وصف العمال الفنية ، والذين كان جل اهتمامهم منصرفًا نحو وصف الأعمال الفنية الأصلية خدمة للأغراض البحثية ، كما كان مسؤولو المصادر البصرية العاملون في الأقسام الأكاديمية أو في المكتبات المتخصصة ، وكذا المسؤولون عن مجموعات الشرائح أو غيرها من البدائل منصرفين نحو وصف الأعمال الفنية والمعمارية للاستخدامات التعليمية داخل الفصول الدراسية . وغالبًا ما كانت هذه المواد توصف في ملفات أو في قواعد بيانات محلية دون تضمينها الفهارس المباشرة للمكتبات ؛ لأن هذه المجموعات قد لا تكون تبعيتها لإدارة المكتبة ، أو لأن تطبيق القواعد الأنجلو أمريكية على فهرستها تكتنفه إشكالات كثيرة.

وكان أخصائيو المصادر البصرية مدركين تمامًا أنهم مضطرون دائمًا إلى وصف شكلين على الأقل من كل عمل ، أولهما العمل الفني أو المعماري الأصلي (والذي يحتمل ألا يكون موجودًا ضمن المجموعة المحلية)، وثانيهما الشريحة التي تمثل ذلك المصدر أو غيرها من البدائل الأخرى للعمل . (وبالطبع كان هناك في معظم الحالات عدد كبير من النسخ يجب التعامل معها، لأن الشرائح ذاتها عادة ما كانت تنتج عن صور فوتوغرافية، هذا فضلاً عن الأعداد الكثيرة من الصور الرقمية المنتجة في صيغ مختلفة والتي كانت تلحق بالشرائح) . وقد بدا تطلع هؤلاء الأخصائيين واضحًا نحو تقاسم التسجيلات التي تصف الأعمال الأصلية مع

غيرهم عن قناعة منهم بعدم جدوى إعادة فهرسة الأعمال الفنية نفسها أكثر من مرة ، كما تطلعوا كذلك إلى وجود خطة ما وراء بيانات تعينهم على وصف الأعمال الفنية الأصلية وتمثيلاتها وصفًا مكتملاً، وخاصة بعد أن اتضحت الصورة بشأن عناصر الوصف التي تخص العمل الأصلي من جهة، وعناصر الوصف التي تخص تمثيلات هذا العمل الأصلي من جهة أخرى.

وقد صيغت الفئات الأساسية لجمعية المصادر البصرية عن قصد في نموذج دبلن كور على اعتبار أنها المجموعة الأساسية من العناصر التي يمكن أن تتقاسمها جميع التطبيقات ، مع ما يلحق بها من عناصر إضافية على المستوى المحلي . وكما هو الحال مع دبلن كور ، لم تحدد لهذه المجموعة تركيبية مسبقة ، ورؤي أن تكون العناصر فيها اختيارية ومتكررة (مع عدم الإقرار الصريح بذلك). وقد اشتملت الإصدارة الأولى من الفئات الأساسية لجمعية المصادر البصرية، على إحدى وعشرين فئة ، توزعت في ثلاث مجموعات تصف إحداها العمل الأصلي (الذي سمي فيما بعد بالكيان object) ، وتصف الثانية المنشئ، وتصف الثالثة التمثيل (الذي سمي بعد ذلك بالبديل surrogate) ، غير أن هذا التوجه روجع مراجعة كاملة في الإصدارة الثانية - " رقم ٢,٠ " - حيث صنفت الفئات في مجموعتين فقط من عناصر البيانات: تختص تسعة عشر منها بوصف العمل الفني نفسه ، وتختص تسعة منها بالتمثيل (الذي يعرف الآن " بالوثيقة البصرية visual document ") . كما أعيد إدراج العناصر الخاصة باسم المنشئ ودوره إلى فئات العناصر الخاصة بالعمل الأصلي ، وحذفت بقية العناصر الأخرى

الخاصة بالمنشئ التي تضمنتها الإصدارة الأولى بعد أن ثبت عدم جدوى تكرارها مع كل عمل ، ورؤي أن يكون من الأحرى وضعها في ملف استنادي مستقل . (لعل مما يلفت الانتباه أنه لم يتح في الإصدارة الثانية إمكانية تقديم معلومات عن المنشئ عند وصف الوثيقة البصرية) .

أما في الإصدارة الحالية من الفئات الأساسية لجمعية المصادر البصرية، الإصدارة " رقم ٣,٠ " ، فلم تعد تقسم الفئات إلى مجموعتين من العناصر المستقلة، التي تخص إحداها العمل الأصلي ، وتخص الأخرى تمثيل هذا العمل (الذي أصبح في سياق هذه الإصدارة الحالية يعرف " بالصورة image ") ، وإنما جمعت كلها في مجموعة واحدة تتألف من سبع عشرة فئة ، تطبق على أي من العمل الأصلي أو صورته ، وقصد من ذلك أن تنشأ تسجيلات مستقلة (أي مجموعات من عناصر ما وراء البيانات) لكل عمل وتمثيله ، ومن ثم أضيفت فئة جديدة تسمى " نوع التسجيلة record type ، تبين ما إذا كانت التسجيلة تخص العمل الأصلي، أم أنها تخص الصورة . وقد جاءت الفئات الأخرى على النحو التالي : النوع Type (أي نوع جنس العمل أو الصورة اعتمادًا على مكنز الفنون والعمارة)، والعنوان، والحجم أو المقاسات، والمادة ، والأسلوب، والتاريخ، والموقع، ورقم الهوية، والشكل، والفترة الزمنية، والخلفية الثقافية، والموضوع، والعلاقة، والوصف، والمصدر، والحقوق .

ولعل مما تجدر الإشارة إليه في هذا السياق ، أن تعريف العمل في سياق "الفئات الأساسية لجمعية المصادر البصرية " يختلف اختلافًا بينًا عن تعريف العمل في سياق المتطلبات الوظيفية للتسجيلات الببليوجرافية FRBR

للاتحاد الدولي لجمعيات المكتبات ومؤسساتها ؛ ففي الوقت الذي يعرف فيه العمل في سياق هذا العمل الأخير - أي المتطلبات الوظيفية للتسجيلة الببليوجرافية - بأنه إبداع فكري أو فني مجرد يتجسد وجوده عندما يقدم في أسلوب تعبير معين وفي وسيلة عرض محددة، يعرف في سياق الفئات الأساسية على أنه " كيان مادي موجود في الحاضر، أو وجد في وقت ما في الماضي، أو سيوجد في المستقبل، وقد يكون هذا الكيان إبداعًا فنيًا كالصور الزيتية أو التماثيل المنحوتة ، أو عملاً استعراضيًا أو ملحنًا أو أدبيًا ، أو مبنى أو نوعًا آخر من أنواع التشييد ، أو قد يكون كيانًا يعكس تراثًا أو ثقافة مادية معينة[١] ". ومن الملاحظ هنا أن العمل في هذا السياق الأخير يجمع ما بين الكيانات التي حددها الاتحاد الدولي لجمعيات المكتبات بدءًا من العمل ، ومروراً بوسيط العرض ، وانتهاءً بمستوى الوحدة التي يتألف منها كيان أكبر.

أما من حيث المفهوم ، فمن الملاحظ أن العمل يقف على طرف نقيض من الصورة ، التي تعرفها الفئات الأساسية لجمعية المصادر البصرية على أنها تمثيل مرئي للعمل ، وقصد بها أن تشمل الشرائح ، والصور الفوتوغرافية ، والملفات الرقمية ، مع ملاحظة أن الصورة نفسها قد تكون عملاً بكل ما تعنيه الكلمة ، ولعل خير مثال على ذلك تلك الصورة التي يصورها أحد المصورين المحترفين لأحد المباني .

---

(١) VRA Core Categories, version ٣,٠, Introduction, available at http://www.vraweb.org/ vracore٣.Htm. Accessed ٢٥ June ٢٠٠٢.

وفي سياق هذه الإصدارة الثالثة "رقم ٣,٠" من الفئات الأساسية لجمعية المصادر البصرية ، يتم وصف كل عنصر بيانات من خلال تسجيل كل من : اسمه، وتعريفه، ومقيداته المعرفة، والقوائم الاستنادية أو اللغات المقيدة المعتمدة لوصف محتواه، والروابط التي تربطه بالإصدارة الثانية من الفئات الأساسية، وبفئات وصف الأعمال الفنية ، وبدبلن كور . ويعرض الشكل رقم (١٠ / ٣) وصفًا لعنصر العنوان بناءً على تلك الإصدارة الثالثة من مواصفة الفئات الأساسية لجمعية المصادر البصرية ، ومع ذلك يلاحظ في هذا الصدد أنه برغم حث المطبقين لهذه المواصفة على تسجيل اسم القائمة الاستنادية المعتمدة لوصف المحتوى ، فإن الخطة لا تتضمن أية وسيلة دلالية تستخدم للإشارة إلى القائمة الاستنادية المستخدمة .

TITLE
Qualifiers:
Title.Variant
Title.Translation
Title.Series
Title.Larger Entity
Definition: The title or identifying phrase given to a Work or an image. For complex works or series the title may refer to a discrete unit within the larger entity (a print from a series, a panel from a fresco cycle, a building within a temple complex) or may identify only the larger entity itself. A record for a part of a larger unit should include both the title for the part and the title for the larger entity. For an image record this category describes the specific view of the depicted Work.
Data Values: formulated according to data content rules for titles of works of art
VRA Core 2.0: W2 Title; V7 Visual Document View Description
CDWA: Titles or Names-Text; Related Visual Documentation-View; Related Visual Documentation-View- Indexing Terms
Dublin Core: TITLE

فئ :

Visual Resources Association, Data Standards Committee. Core Categories for Visual Resources, version ٣,٠. Available at http://www.vraweb.org/vracore٣.htm.

ومن الملاحظ أن فكرتي مقيدات العنصر والترميز النقطي "dot" notation لتمثيل هذه المقيدات اقتبستا من خطة دبلن كور ، واعتبرت مقيدات لتوضيح العناصر ، وتحديد معنى الفئة ؛ ومن ثم تأتي مقيدات عنصر " التاريخ Date " مثلاً هكذا على النحو التالي :

Date. Creation

Date.Design

Date.Begining

Date.Completion

Date.Alteration

Date.Restoration

ومع ذلك يأتي نموذج البيانات بالنسبة لتقييد العناصر في صيغة أكثر مرونة من الصيغة التي يتخذها في دبلن كور، حيث يسمح في سياقه بالتركيبات التالية على سبيل المثال:

Creator.Attribution

Creator.Personal name

Creator.Corporate name

ومما يلاحظ كذلك أن عنصري الدور والخاصية قد يخفقا في اختبار "التهميش أو التجاهل dumb down"، كما أن معالجة عنصر الدور (هكذا Creator.Role) على أنه أحد العناصر وليس على أنه مقيد حقيقي (أي Creator.Artist)، تعني أن أسماء المنشئين وأدوارهم يجب أن يأتيا دائمًا

مقترنين، بالرغم من عدم وجود آليات لربط عناصر البيانات ذات الصلة بعضها مع بعض داخل التسجيلة .

ومع أن وجود آلية للربط بين تسجيلات الفئات الأساسية لجمعية المصادر البصرية ينظر إليها البعض على أنها من قضايا التطبيق ذات الطابع الداخلي التي تعالج خارج مجال خطة ما وراء البيانات الوصفية نفسها، فإنه من المفترض وجود مثل هذه الآلية، ومن الضروري الربط بين التسجيلات الخاصة بالأعمال الأصلية من ناحية، والتسجيلات الخاصة بصور هذه الأعمال من ناحية أخرى، كما يفترض كذلك أن يحترم مستخدمو الفئات الأساسية مبدأ "واحدة لواحد ١:١ "، أي أن يخصص للعمل الواحد تسجيلة واحدة، فإذا أجري مسح ضوئي لأحد الرسوم، ومن ثم نتج عن ذلك صورة رقمية في صيغة (جي بيج) مثلاً، فينبغي في هذه الحالة أن تخصص تسجيلة لوصف الرسم الأصلي، وتسجيلة أخرى لوصف تلك الصورة الرقمية، وسوف تبدو التسجيلتان وكأنهما تسجيلتان مختلفتان من حيث المنشئ، ومن حيث التفاصيل المادية كالحجم والوسيط، مع ما قد يكون بينها من اختلاف أو اتفاق في العنوان والموضوع والخلفية الثقافية .

إن تلك التغيرات التي طرأت على الفئات الأساسية لجمعية المصادر البصرية من إصدارتها الأولى إلى إصدارتها الثالثة تعكس وعيًا متناميًا بتعقد كم العلاقات التي ينبغي إثباتها في هذا الصدد، وليس هذا فحسب بل بتعقد أنواع هذه العلاقات ؛ حيث يمكن أن يكون لكيان فني واحد عدة تمثيلات محملة على وسائط متنوعة (مثل الشرائح، والصور الفوتوغرافية ، والصور الرقمية)، بل قد تختزن تلك الوسائط في صيغ رقمية مختلفة (

مثل صيغة تيف TIFF، أو صيغة جي بيج JPEG، أو صيغة جيف GIF)؛ كما أن تلك التمثيلات قد تصور الكيان بأكمله، أو جزءًا منه، أو مجرد لقطات منه. وقد تتوافر للتمثيلات خصائص لا تتوافر للكيان نفسه، كما هو الحال بالنسبة لخصائص الإضاءة التي ترتبط بالصورة وليس بالعمل الأصلي ، كما يمكن أن تكون الأعمال مرتبطة بعضها ببعض، كما هو الحال بالنسبة لعلاقة العمل الفردي بالمجموعة group: item ، وعلاقة الجزء بالكل whole: part ، وما شابه ذلك من العلاقات الاشتقاقية .

ويمكن في سياق الإصدارة الثالثة من الفئات الأساسية إنشاء عدد لا حدود له من تسجيلات الصور، وربطها بتسجيلات الأعمال الأصلية التي تمثلها ، ومع إمكانية ربط الأعمال الأصلية وصورها فنيًا من خلال بعض تطبيقات قواعد البيانات ، يمكن ربط الأعمال المتصلة بعضها ببعض دلاليًا بطريقتين ؛ حيث يفترض من ناحية في حالة علاقة الجزء بالكل whole: part أن يعطي اسم الجزء في العنصر المسمى "Title.Larger Entit "، كما أن هناك من ناحية أخرى أشكالاً أخرى من العلاقات ينبغي وصفها عن طريق استخدام الفئة المسماة "العلاقة Relation " - تلك الفئة التي توصف في مواصفة هذه الإصدارة الثالثة على أنها تأتي في شكل عنصرين ثنائيين ، هكذا :

Relation.Identity

Relation.Type

مع أنها موضحة في الأمثلة على أنها عنصر مقيد وحيد ، هكذا كما يتضح من المثال التالي :

Relation.derived from = Drawing by Gerog Pencz in the staatsarchiv, Nuremberg, Germany

وإنني أشعر وأنا أتعامل مع هاتين الإصدارتين الأخيرتين(١) من الفئات الأساسية لجمعية المصادر البصرية أنني أمام عمل لا يزال في طور الإعداد، ولمّا يكتمل بعد، عمل مليء بالأخطاء الطباعية، والتعريفات الهزيلة، والأمثلة المتضاربة، ومع كل ذلك فقد حظيت هاتان الإصدارتان بقبول سريع وانتشار واسع بسبب إمكانية إدارة تلك الفئات، مع تعددها وتعدد مقيداتها، ولكن بشكل أفضل بكثير من إمكانية إدارة فئات ائتلاف وصف الأعمال الفنية CDWA، ناهيك عن أن عناصر البيانات في هاتين الإصدارتين جاءت متناغمة بشكل جيد مع مجموعات المصادر البصرية . وجدير بالإشارة هنا إلى أن جمعية المصادر البصرية تعمل الآن في مسودة دليل بعنوان " فهرسة الكيانات الثقافية Cataloguing Cultural Objects " ، يتوقع له أن يضيء الطريق أمام التطبيقات، ويحل كثيرًا من الإشكالات التي تفاقمت في ظل غياب التوثيق .

وإذا كانت ثمة توجيهات واضحة لاستخدام بعض المكتبات للفئات الأساسية لجمعية المصادر البصرية ، فلعل ذلك جاء من منطلق الافتراض بأن وجود مجموعة عناصر أساسية تشجع على التناغم وعلى توافق العمل البيني بين التطبيقات المختلفة(٢) . ولعل من أبرز تطبيقات الفئات الأساسية

(١) تقصد الثانية والثالثة [المترجم] .

(١) Committee on Cataloging: Description and Access, Task Force on VRA Core Categories, Summary report, ٢٠٠١, available at http://www.ala.org/aicts/organization/ccs/ccda/tf-vral.html#repor. Accessed ٢٥ June ٢٠٠٢.

لجمعية المصادر المرئية والتي يمكن أن يشار إليها في هذا الصدد " فهرس جامعة هارفارد لمعلومات المواد البصرية Harvard University's VIA Catalog ، والمستودع التعاوني للصور الأكاديمية Academic Image Cooperative " ، وفهرس إدارة بيانات الفنون البصرية the Visual Arts Data Service Catalog .

وقد استخدمت جامعة هارفارد الإصدارة الثانية من الفئات الأساسية مع بعض التعديلات اللازمة كخطة أساسية لما وراء البيانات لفهرسها الخاص بإتاحة المعلومات البصرية(١)؛ حيث يتيح هذا الفهرس إمكانات الوصول الموحد إلى المجموعات البصرية المتوافرة في عشرة مستودعات مختلفة بجامعة هارفارد ، متضمنًا بذلك التسجيلات المترابطة للمجموعات (أي مجموعات المواد ذات الصلة) ، والأعمال (الأعمال الفنية المنفردة) ، والبدائل (الصور والتمثيلات الفردية) . كما طبقت إدارة بيانات الفنون البصرية التي تعد جزءًا من إدارة بيانات الفنون والإنسانيات في المملكة المتحدة U.K.'s Arts and Humanities Data Service الإصدارة الثالثة من الفئات الأساسية بهدف تقديم خدمات استشارية في مجال الأرشفة الرقمية(٢)، ويشتمل فهرس هذه الإدارة على توصيفات وصور مصغرة

---

(٢) Harvard University Library, Visual Information Access, available at http://via.harvard.edu:٧٤٨/htmI/VIA.html. Accessed ٢٥ June ٢٠٠٢.

(١) Visual Arts Data Service (home page), available at http://vads.ahds.ac.uk/index.html. Accessed ٢٥ June ٢٠٠٢.

لمواد بصرية من عشر مجموعات مختلفة ، تتراوح ما بين مجموعات المنسوجات إلى مجموعة الفنون المعمارية . وقد دفعت هذه الإدارة إلى تطبيق تلك الإصدارة ثلاثة عوامل ، هي : الشهرة التي اكتسبتها فئات وصف المصادر المرئية من ناحية ، وتشجيعها الاعتماد على اللغات المقيدة من ناحية ثانية ، وتوفيرها إمكانية الربط بين ائتلاف وصف الأعمال الفنية و دبلن كور من ناحية ثالثة(١).

أما المستودع التعاوني للصور الأكاديمية ، فكان مشروعًا تخطيطيًا ونموذجًا أوليًا لاتحاد المكتبات الرقمية بتمويل من مؤسسة ميلون(٢). وكان الهدف من وراء إنشائه تطوير قاعدة بيانات للصور الفنية تدعم برامج تدريس تاريخ الفنون، ولهذا الغرض طور نموذج ما وراء بيانات يتوافق مع الإصدارة الثالثة للفئات الأساسية لوصف الصور التي سيشملها المشروع وصفًا كاملاً ، وفي نهاية المطاف أدمجت قاعدة بيانات هذا المشروع مع مشروع أكبر عرف بمبادرة مستودع الفنون أو اختصاراً آرت ستور(٣) . ArtSTOR

---

(٢) Phil Purdy. "Digital Image Archiving and Advice: In Tandem with the Visual Arts Data Service (VADS)," Cultivate Interactive ٤ (May ٢٠٠١), available at

http://www.cultivate-int.org/issue٤/vads. Accessed ٢٥ June ٢٠٠٢.

(٣) Digital Library Federation, Academic Image Cooperative, available at

http://www.diglib.org/collections/aic.htm. Accessed ٢٥ June ٢٠٠٢.

(١) مشروع تبنته مؤسسة ميلون لتطوير الصور الرقمية وتخزينها وتوزيعها وما يتصل بها من مواد بحثية إلكترونياً بهدف خدمة أغراض دراسة تاريخ الفنون والعمارة وغيرهما من مجالات العلوم الإنسانية [المترجم].

READINGS :

The CDWA and the VRA Core specification are available online:

- Baca, Murtha. "A Picture Is Worth a Thousand Words: Metadata for Art Objects and Their Visual surrogates." In Wayne Jones et al., eds., Cataloging the Web: Metadata, AACR, and MARC٢١. Lanham, Md.: Scarecrow, ٢٠٠١.
- Baca, Murtha, and Patricia Harpring, eds. Categories for the Description of Works of Art. The J. Paul Getty Trust and College Art Association, ٢٠٠٠. Available at http://wiww.getty.edu/research/institute/standards/cdwa/.
- Visual Resources Association, Data Standards Committee. Core Categories for Visual Resources, version ٣,٠. Available at http://www.vraweb.org/vracore٣.htm.
- Visual Materials: Processing and Cataloging Bibliography. Prints and Photographs Division, Library of Congress, Washington, D.C. Available at http://www.loc.gov/rr/print/vrnbib.html.
- Special issues of two journals focus on the CDWA and the VRA Core, respectively:
- Baca, Murtha, and Patricia Harpring, eds. "Art Information Task Force Categories for the Description of Works of Art." Visual Resources ١١, no. ٣/٤ (١٩٩٦), special issue.

- "The VRA Core Categories." VRA Bulletin ٢٥, no. ٤ (winter ١٩٩٨), special issue.
- Museum Information Standards (web page). Available at http://www.diffuse.org/museums.html. Accessed ٢٦ August ٢٠٠٢.
- The European Commission's Diffuse Project (http://www.diffuse.org) documents standards and specifications facilitating information exchange. Their web pages include a comprehensive list of museum information standards.

الفصل الحادي عشر

جيلس والمعلومات الحكومية

تشير التسمية الاستهلالية جيلس GILS إلى عدة أشياء ؛ ففي أضيق معانيها تشير إلى " خدمة محدد موقع المعلومات الحكومية Government Information Locator Service "، وهي مبادرة فدرالية يجري تطبيقها على الوزارات والهيئات التي تشكل القطاع التنفيذي من الحكومة(١) فقط . كما تعرف المواصفات الفنية لتنفيذ هذه المبادرة والتي تشتمل على مجموعة أساسية من عناصر ما وراء البيانات " ببروفايل جيلس GILS profile "، ولكن يشار إليها تجاوزًا بجيلس GILS. أما في أوسع معانيها ، فهي تشير إلى أي تطبيق يستخدم بروفايلات جيلس، حيث وجدت أكثر من مبادرة حكومية لتنفيذ جيلس . وفي معنى أعم من كل ذلك، أدى تبني بروفايل جيلس خارج حدود الولايات المتحدة الأمريكية من ناحية، وفي مجال المعلومات غير الحكومية، إلى وضع مصطلح آخر، وهو" خدمة محدد موقع المعلومات الكونية أو الشاملة Global Information Locator Service "، ليشمل جميع تلك الاستخدامات الأوسع .

ولبرنامج جيلس جذوره في سياسة المعلومات في كل من إدارة كلينتون ومبادرة البنية التحتية للمعلومات الوطنية National Information Infrastructure (NII) initiative. وقد صدر التكليف الرسمي لهذا البرنامج في عام ١٩٩٤م، وأشير إليه صراحة في نشرة مكتب الإدارة والموازنة OMB Bulletin ٩٥-٠١ هكذا : " إنشاء خدمة محدد موقع المعلومات الحكومية " (Office of management and Budget, ١٩٩٤). وتجسدت الرؤية الأساسية لجليس في أن يكون بمثابة فيدرالية من " خدمات محددات مواقع الهيئات القابلة للعمل

(١) تقصد بالطبع حكومة الولايات المتحدة [المترجم].

البينـي " تعمل على توفير سبل الوصول إلى مصادر المعلومات التي تتيحها تلك الهيئات . وتضمنت النشرة توصيفًا لمسؤوليات الهيئات تجاه تنفيذ جليس ، وتحديـد الأهـداف المنـشودة مـن التعـريف بمـصادر المعلـومات العامـة ، ووصـف المعلـومات التـي تتضمنها تلك المصادر ، وتقديم العون للحصول علـى تلك المعلـومات ، هذا فضلا عن الاتقاء بالإدارة الالكترونية لسجلات تلك الهيئات . وبنـاء علـى ذلك يمكن القول إن جيلس يسعى لتحقيق هدفين : أولهمـا عـام يتمـثل فـي الارتقـاء بـسبل إتاحـة الوصـول إلـى المعلـومات ، وثانـيهما محلـي يتمـثل فـي تحسين أداء إدارات الأرشـيف الجـاري record management في داخل الهيئات .

وتعميمًا للمشروع نـشرت المواصـفات الخاصـة بتنفيذه في نشرة معايير معالجـة المعلـومات الفيدرالية FIPS pub ، رقم ١٩٢ ، وجـاءت بعنوان : بــروفايل التطبــيق الخــاص بخدمــة محــدد موقــع المعلــومات الحكومــية (Application profile for the Government Information Locator Service ، National Institute for Standard and Technology, ١٩٩٤).

ومـع أن معظم خطط ما وراء البيانات عادة ما تعرف مستقلة عن خدمة الـبحث التـي تـستخدم للوصـول إلـيها ، فقد كـتب بـروفايل جـيلس كبـروفايل تطبـيق لبـروتوكول الـبحث Z٣٩,٥٠ ، بـل وفـضلاً عـن تحديـده لمجمـوعة العناصـر الأساسـية، يـتطلب بـروفايل جـيلس مـن البـرمجية الخادمـة لجـيلس GILS server أن تعمـل كبـرمجة خادمـة للبـروتوكول Z٣٩,٥٠ ، كمـا يطلـب مـنه كذلك أن يحدد بشكل دقيق كلاً من مجموعة الخصـائص attribute set ، والمجمـوعة التشخيصية diagnostic set ، وغيرهما من الخصائص الأخرى

التي يدعمها - أي البروتوكول Z٣٩,٥٠ - حتى يمكن له أن يصبح متوافقًا مع جيلس GILS-compliant .

وتضمن الملحق الخامس - رقم E - من بروفايل جيلس تعريفًا بالعناصر الأساسية لجيلس ، والخصائص المحددة لهذه العناصر ، والتي جاءت في ثلاث خصائص هي : اسم العنصر ، ومدى قابلية هذا العنصر للتكرار ، ثم تعريفًا له مع ملاحظة أن جميع العناصر اختيارية . وبالرغم من عدم اشتمال المواصفة على قواعد للمحتوى ، فإن تعريف بعض العناصر تطلب أن تسجل القيم الخاصة به في صيغة معينة ، أو اعتمادًا على لغة تكشيف مقيدة . ويتضمن الشكل رقم (١١ / ١) العناصر السبعة الأولى المعرفة في الملحق المشار إليه سابقاً .

وتُعنى تسجيلات ما وراء بيانات جيلس ، والتي تسمى تسجيلات محدد الموقع locator records ، بوصف جميع أشكال مصادر المعلومات الخاصة بالهيئات ، ومن ثم فهي لا تقتصر على المطبوعات الفردية وحدها ، وإنما تتسع لتشمل قواعد البيانات، والفهارس ، والأدلة ، والخدمات المتاحة على الخط المباشر، ومواقع الإنترنت، بل تتسع بشكل أكبر لتغطي المصادر غير الببليوجرافية كقوائم المهام Joblines والبرامج ؛ ولهذا السبب، ومع وجود العناصر الأساسية للوصف الببليوجرافي (أي عناصر: العنوان، والمنشئ، والمشارك ، و تاريخ النشر، ومكان النشر، ولغة المصدر، والمستخلص)، فقد صيغت التعريفات بشكل موسع حتى تطبق على وصف أكبر عدد من مصادر المعلومات.

Title (Not Repeatable) This element conveys the most significant aspects of the referenced resource and is intended for initial presentation to users independently of other elements. It should provide sufficient information to allow users to make an initial decision on likely relevance. It should convey the most significant information available, including the general topic area, as well as a specific reference to the subject.

Originator (Repeatable) This element identifies the information resource originator.

Contributor (Repeatable) This element is used if there are names associated with the resource in addition to the Orginator, such as personal author, corporate author, co-author, or a conference or meeting name.

Date Of Publication (Not Repeatable) The discrete creation date in which the described resource was published or updated, though not for use on resources that are published continuously such as dynamic databases. Date of Publication Textual may also provide additional information such as when the resource was originally published. This element may be expressed in one of two forms:

- **Date Of Publication Structured** Date described using the ISO 8601 prescribed structure (fixed 8 characters, YYYYMMDD).
- **Date Of Publication Textual** Date described textually.

Place of Publication (Not Repeatable) The city or town where the described resource was published. May also include country if location of city is not well known.

Language of Resource (Repeatable) This element indicates the language(s) of the described resource as represented by the MARC three character alpha code. If a resource is multilingual, repeat this element for each applicable language.

Abstract (Not Repeatable) This element presents a narrative description of the information resource. This narrative should provide enough general information to allow the user to determine if the information resource has sufficient potential to warrant contacting the provider for further information.

آخرين، هما: الغرض، ويهدف إلى وصف مبررات إنشاء أو إتاحة مصدر

المعلومات، والبرنامج، ويشير إلى برنامج أو برامج الهيئة التي يدعمها المصدر. ومما تجدر ملاحظته في هذا الصدد ذلكم التشابه الواضح بين الفئات الأساسية لجيلس من ناحية، وصيغة مارك ٢١ لمعلومات المجتمع المصممة لوصف المصادر غير البيليوجرافية كالمنظمات والبرامج والخدمات التي يفيد منها المجتمع من ناحية ثانية.

ويستخدم عنصر الإحالة cross reference ، القابل للتكرار لتجميع العناصر الفرعية التي تُعرف بالتسجيلات الأخرى لمحدد الموقع ، أو بمصادر المعلومات الأخرى التي لها علاقة معينة بالمصدر الذي يتم وصفه . كما تشتمل الفئات الأساسية لجيلس على عناصر أخرى تعنى بوصف التغطية الجغرافية والزمنية للمصدر؛ حيث تعرف العناصر الفرعية لعنصر المجال المكاني(١) Spatial Domain بخطوط الطول والعرض لإحداثيات الحدود الشمالية والجنوبية والشرقية والغربية للمكان ، هذا في الوقت الذي تسجل فيه الأسماء الجغرافية في العنصر الفرعي " المكان Place " ؛ ذلك العنصر المعرف - كما هو الحال بالنسبة للكشاف الموضوعي المقيد - بعنصر فرعي خاص بالمكنز ، وعنصر فرعي آخر خاص بالمصطلح أو المصطلحات ، هكذا :

المكان (عنصر متكرر) : يعرف هذا العنصر الفرعي المواقع

---

(١) يقصد بالمعلومات المكانية spatial information وأحياناً تعرف Geospatial Information : المعلومات المرجعية للأماكن الجغرافية مثل خطوط الطول والعرض وما شابهها ، وذلك غير معلومات المكان أو المنطقة الجغرافية place التي تركز عادة على الموقع الجغرافي [المترجم].

الجغرافية التي تحددها مجموعة البيانات أو مصدر المعلومات ، وذلك عن طريق تركيبتين مرتبطتين هما:

مكنز الكلمات الأساسية للأماكن (عنصر غير متكرر) : اسم أحد المكانز المسجلة رسميًا ، أو أحد المصادر المعتمدة المشابهة الخاصة بالكلمات الأساسية لأسماء الأماكن . على أن تعطي كل كلمة أساسية في الحقل التابع المتكرر ، وهو :

الكلمة الأساسية للمكان (عنصر متكرر) : الاسم الجغرافي لأحد المواقع الذي تغطيه مجموعة البيانات أو مصدر المعلومات .

وتجدر الإشارة في هذا الصدد إلى أن مواصفة جيلس لا تتضمن قواعد المحتوى لفئاته الأساسية ، وإن كانت تتضمن بعض التعليمات القليلة التي ترد في ثنايا تعريفات بعض العناصر ، ولكن هناك دعوات مشجعة لاستخدام قواعد محتوى خارجية ، لعل من أكثرها تداولاً " إرشادات إعداد مداخل الفئات الأساسية لجيلس Guidelines for the preparation of GILS Core Entries " التي أعدتها إدارة الأرشيفات والمحفوظات الوطنية(١) the National Archives and Records Administration (NARA) ، حيث طرحت هذه الإرشادات توجيهات إضافية لكيفية صياغة قيم العناصر ، وحددت الحالات التي يعامل فيها العنصر على أنه عنصر اختياري ، والحالات التي

(١) National Archives and Records Administration, Guidelines for the Preparation of GILS Core Entries, published online by the Defense Technical Information Center, available at http://www.dtic.mil/gils/documents/naradoc/ Accessed ٢٦ June ٢٠٠٢.

يعامل فيها على أنه عنصر إجباري ، كما أعطت أمثلة على الاستخدام الصحيح لمعظم العناصر ، وأكدت كذلك على استخدام اللغات المقيدة بالنسبة للأسماء والمصطلحات الكشفية ، ومما أشارت إلى وجوب استخدامه في هذا الصدد وتحديدًا بالنسبة لأسماء الهيئات " دليل حكومة الولايات المتحدة U.S. Government Manual " ، كذلك أوصت هذه الإرشادات باستخدام عنصر " الإحالة " لعمل روابط إلى المكانز المستشهد بها والمتاحة على الخط المباشر ، على أن تتم تلك الإشارة في العنصر الفرعي " مكنز " التابع لعنصر " اللغة المقيدة " . وأخيرًا تجدر الإشارة إلى أن هذه المواصفة يرافقها ملحق بنماذج الممارسات الجيدة للمداخل الأساسية التي تصف مختلف أنواع الكيانات . وفي الشكل رقم (١١ / ٢) مثال لتسجيلة نموذجية لفئات جيلس الأساسية .

وفي عام ١٩٩٧م نشر تقرير بني على إحدى الدراسات التقييمية التي أعدها مجلس جيلس للسنتين الأوليين من البرنامج الفيدرالي بناء على طلب من الأرشيفيين الأمريكيين ، وخلص التقرير إلى نتيجة مفادها أنه بالرغم من استمرار صلاحية الرؤية التي كان البرنامج قد بني عليها والمتمثلة في مساعدة المستفيدين على تحديد مواقع المعلومات الحكومية المتاحة للعامة ، فإن الهدف المنشود من ورائه والمتمثل في إنشاء خدمة محدد لمواقع المعلومات الحكومية جميعها لم يتحقق بعد، وبدلاً من تحقق هذا الهدف فقد أثبتت الدراسة التقييمية أن تطبيقات الهيئات لجيلس لم تكن متساوية ، ولم تأت متناغمة ، بل جاءت مستقلة بعضها عن بعض، يضاف إلى ذلك تزايد عدم وضوح الرؤية حول الأدوار النسبية التي يقوم بها كل من جيلس

ومواقع الهيئات على الوب ، وقد أوصى التقرير في نهايته بأن يعيد برنامج جيلس تركيزه على مهمة التعريف بالمعلومات الحكومية المتاحة إلكترونيًا دون غيرها من المعلومات ، مع عمل الروابط اللازمة لها ، وأن يتخلى عن وظيفة إدارة الأرشيفات الجارية ، كما أوصى بإعادة التأكيد على توسعة مجال هذه الخدمة لتكون على مستوى جميع الهيئات الحكومية . وإذا كان هذا هو التقرير فليس من الواضح ما ستثمر عنه من نتائج ، وإن كانت هناك تحسينات تجرى على عمليات بحث جيلس ؛ حيث يلاحظ أن خدمة محدد مواقع المعلومات الحكومية التي يديرها مكتب المطبوعات الحكومية GPO تقوم بتجميع قواعد البيانات من خمس وثلاثين هيئة فيدرالية مختلفة ، منها قاعدة بيانات مكتب المطبوعات الحكومية نفسه ، كما تقوم كذلك بتقديم خدمة البحث الفيدرالي لموقعين آخرين تجميعيين من مواقع جيلس(١) .

Automated Information System Example

Title: Retained Records Database

Acronym: RET

Originator:

Department/Agency Name: National Archives and Records Administration
Name of Unit: Office of Records Administration

Local Subject Index:

Local Subject Term: US Federal GILS

Abstract: The Retained Records Database (RET) contains descriptions of unscheduled records as well as scheduled records that other Federal agencies have not transferred to the National Archives. Information contained in the database also includes the conditions under which the records are maintained, where they are located, the contact person, a tickler date indicating when the records should be re- evaluated, and scheduling data. The History File contains records previously listed in RET that have now been transferred to the National Archives or a Federal Records Center. The Oral History File contains descriptions of oral history projects through 1992. Routine updates and additions to the system occur semi- annually.

Begin Date: 1990

Purpose: The Retained Records Database was created to provide a centralized source of information about older series of permanent or potentially permanent records maintained in agency custody so that these records can be tracked and eventually transferred to the National Archives.

Agency Program: 44 U.S.C. 29 authorizes the Archivist of the United States to undertake certain records management functions.

Time Period of Content: 1755 -

Availability:

Distribution:
Name: Office of Records Administration
Organization: National Archives and Records Administration
Street Address: 8601 Adelphi Road

(١) U.S. C
Inform
http://

أسا

الشكل رقم (٢/١١) بداية لعينة إحدى تسجيلات فئات جيلس الأساسية GILS CORE . المصدر:
National Archives and Records Administration, Guidelines for the Preparation of GILS Core Entries, published online by the Defense Technical Information Center,

وقد تبنت عدة ولايات استخدام جيلس في تقديم خدمات إتاحة المعلومات الحكومية الخاصة بها، حيث يشجع برنامج البحث الموسوم "Find-it" في كل من ولاية إلينوي وواشنطن وبعض الولايات الأخرى، الهيئات التابعة لتلك الولايات على تضمين ما وراء البيانات الأساسية لجيلس في صفحات مواقعها الخاصة بالوثائق الإلكترونية العامة، على أن تقوم برامج التكشيف العنكبوتية التي تديرها كل ولاية بعد ذلك بجمع ما وراء البيانات وإتاحة إمكانات الوصول إليها للمستفيدين عبر خدمات بحث مركزية . وفي هذا الصدد تقدم بعض الولايات برامج تدريبية على تطبيق جيلس ، في حين تلجأ بعض المواقع الأخرى إلى إتاحة نماذج بسيطة على الوب لإدخال ما وراء البيانات التي لا تظهر تفاصيل مجموعة عناصر الفئات الأساسية لجيلس . ولأن قليلاً من مبادرات المعلومات الحكومية هي التي فضلت استخدام

مجموعة عناصر ما وراء بيانات أخرى مثل دبلن كور على جيلس، وبحكم تفاوت الإرشادات والممارسات من ولاية لأخرى ، فقد بذلت بعض الجهود المشجعة للبحث البيني بين تطبيقات جيلس على مستوى الولايات cross-state searching . وفي سبيل تحسين عمليات البحث بين الولايات عن طريق الاعتماد على لغة مقننة موضوعية متجانسة، طورت - على نحو تعاوني- قائمة استناد موضوعية عامة ، تعرف " بشجرة موضوعات جيلس GILS Topic Tree ".

ولتجاوز إشكالية التباين بين عناصر ما وراء البيانات ، قامت كل من ولاية واشنطن وأوتاوا ومينسوتا بتجربة إجراء بحث بيني عبر الولايات باستخدام أسلوب المعرفات الرقمية numeric identifiers الذي يعرف برموز زد Z tokens، وهي عبارة عن معرفات رقمية تربط كل عنصر من عناصر ما وراء البيانات بخاصية معينة من خصائص المجموعة المسماة بالخصائص الببليوجرافية الأولى لزد٣٩,٥٠ (١) Z٣٩,٥٠ Bib-١ attribute set ؛ حيث يمكن - على سبيل المثال - ربط عنصر " المنشئ " في دبلن كور ، " والمنشئ " في جيلس معًا بالرمز ١٠٠٣ Z token الخاص باسم المؤلف Author-mane ، والذي يمكن استخدامه بعد ذلك في عمليات البحث البيني عبر الولايات .

ومن الملاحظ أن الفئات الأساسية لجيلس لا تستخدم على المستوى

---

(١) Philip Coombs, White Paper on the Use of Numeric Tokens in Resource Descriptions, ٢٠ September ١٩٩٩, available at

http:/www.statelib.wa.gov/projects/imls/tokens.htm. Accessed ٢٧ June ٢٠٠٢.

الدولـي لوصـف المعلـومات الحكومـية ، حـيث رأت خدمـة محـدد المواقـع الحكومـية الأسـترالية[١] Australian Government Locator Service (AGLS) أن تـستخدم مجمـوعة عناصـر أساسـية تعتمد علـى دبلـن كـور ، مـع إجـراء بعض التوسـعات القليلة ، من أهمها : إضافة عنصر " الوظيفة Function " لبيان مجـال عمـل الجهـة الحكومـية ، وعنصـر " مدى التوافر Availability " ليحدد كيفية الحصول علـى مصدر المعلـومات أو الاتصـال بـه ، وعنصـر " الجمهور Audience " للإشـارة إلى الجمهور المستهدف للمصدر ، وعنصر "مدى قانونـية المصـدر Mandate " للإشـارة إلـى السند القانونـي الذي يطلب علـى أساسـه إنـشاء المصـدر أو تقديمـه[٢]. كذلك اسـتخدام دبلـن كـور أساسًـا لخطـط مـا وراء البيانات للمعلومات الحكومية فـي كل من كندا والدنمارك، وفنلندا، وأيرلندا، ونيوزيلندا، والمملكة المتحدة .

وقد أوصـى تقريـر أعدتـه إحدى لجان العمل التي شكلها الفرع الرئيسي الكندي لـضباط المعلـومات للمقارنـة بين جيلس ودبلن كور واختيار أحدهما

---

(٢) Australian Government Locator Service (home page), available at http:/www.govonlin.gov.au/projects/standards/agls.htm. Accessed ٢٧ June ٢٠٠٢. See also the AGLS User Manual at http://www.naa.gov.au/recordkeeping/gov_online/agls/user_manuaVintro.html

(٣) Dublin Core Metadata Initiative, "Adoption of Dublin Core by Governments," available at http://dublincore.ogr/news/adoption/. Accessed ٢٧ June ٢٠٠٢.

أساسًا لمعيار ما وراء البيانات للمعلومات الحكومية ، بمعيار بدبلن كور[١].
ومن الجوانب التي سجلها التقرير كسلبيات لجليس ، تعقده النسبي ، وانخفاض معدلات تبنيه داخل الولايات المتحدة ، وعدم إتاحته الإمكانات الكاملة للمستفيد ليتدخل في تركيبته .

وآخر ما تجدر الإشارة إليه في هذا الصدد أن مجموعة العمل الحكومية في مبادرة دبلن كور لما وراء البيانات قدمت مسودة لأحد الملفات التفصيلية أو بروفايلات التطبيقات لاستخدام دبلن كور في المجال الحكومي ، غير أنه لا يزال قيد المراجعة[٢]، وها نحن في انتظار بيان من مجموعة العمل يوضح العلاقة بين جيلس ودبلن كور .

READINGS :

- Global Information Locator Service (GILS) (home page). Accessed ٢٧ June ٢٠٠٢. Available at http://www.gils.net/.

  The federal GILS home page, with links to information about GILS initiatives, implementations, policy background, tools, and the GILS profile. The GILS Topic Tree is available on the official

---

(١) Government On-Line Ad hoc Interdepartmental Metadata Working Group, "Selecting and Implementing a Metadata Standard for the Government of Canada," ٢٢ March ٢٠٠١, available at http://www.cio-dpi.gc.ca/im-gi/references/meta-standard/meta-standard٠٠-e.asp. Accessed ٢٧ June ٢٠٠٢.

(٢) Dublin Core Metadata Initiative, DCMI Government Working Group (web page), available at http://dublincore.org/groups/govrnment/. Accessed ٢٧ June ٢٠٠٢.

GILS site at http://www.gils.netitrees.html. Information on the history and development of the topic tree is collected at http://www.fidocat.com/gils/ .

- Moen, William E., and Charles R. McClure. An Evaluation of the Federal Government's Implementation of the Government Information Locator Service, Final Report. ٣٠ June ١٩٩٧. Available at http://www.access.gpo.gov/su_docs/gils/gils-eval .

  the evaluation of federal GILS implementations gives a thorough history of the GILS movement from a legislative and policy perspective, as well as describing the evaluation itself and its conclusions and recommendations.

الفصل الثالث عشر

أونيكس الدولـــي

أونيكس الدولي (أو إرشادات تبادل المعلومات المتاحة على الخط المباشر (Guidelines for Online Information eXchange (ONIX) عبارة عن خطة طورها في بداية الأمر مجموعة من الناشرين [الأمريكيين] لتبادل المعلومات التجارية في صيغة إلكترونية فيما بينهم وبين تجار التجزئة للمواد الإلكترونية[1] e-tailer، وتجار التجزئة ، وتجار الجملة ، والموزعين ، وغيرهم من الأطراف الأخرى المرتبطة بعالم تجارة الكتب . وقد كانت هناك - في حقيقة الأمر - خطتان لما وراء البيانات طورتا لأغراض مشابهة على الجانب الآخر من المحيط ؛ حيث صدرت برعاية إحدى المنظمات الدولية التي تدعو لتشجيع التجارة الإلكترونية في قطاعي الكتب والمسلسلات ، والمعروفة باسم إديتير EDItEUR ، مسودة معيار إديتير لتبادل معلومات الإنتاج (أو اختصارًا إيبيكس EPICS)[2] . وقد تبنت تطوير الإصدارة المبدئية من أونيكس جمعية الناشرين الأمريكيين (Association of American Publishers (AAP ، وظهرت في يناير من عام ٢٠٠٠م الإصدارة الأولى منها وهي الإصدارة رقم ١,٠ ، كما ظهرت في الوقت نفسه الإصدارة رقم ٣,٠٢ من معجم بيانات إيبيكس EPICS Data Directory . وبعد فترة وجيزة بدأت جهود توحيد هذين العملين، حيث أطلقت في مايو من العام نفسه إصدارة جديدة من أونيكس أكثر توافقاً مع إيبيكس، وموجهة لخدمة التطبيقات الأوربية والأمريكية على حد سواء، ومن ثم جاءت باسم أونيكس الدولي ١,٠١ ONIX International ١,٠١ ، في الوقت نفسه أعيد تعريف إيبيكس لتصبح معجم بيانات أكثر شمولاً يمكن أن يظهر

(١) e-tailer هي اختصار للجملة electronic retailers [المترجم].

(٢) اختصار لـ EDItEUR Product Information Communication Standards [المترجم].

فيه أونيكس كمجموعة فرعية . وأوكلت رعاية الخطتين لمنظمة إديتير تحت إشراف لجنة توجيه دولية واحدة . وفي الوقت الذي لا تزال فيه آخر إصدارات إيبيكس هي الإصدارة رقم ٣,٢ ، انطلقت جهود توسعة أونيكس والدعاية لتطبيقه، وحدّثت مواصفته الأساسية لتشمل البيانات الخاصة بالكتب الإلكترونية، ولتظهر بالتالي في إصداراتها التالية - الإصدارة رقم ٢,٠ ، والتي تعرف بتسجيلة منتج أونيكس ONIX product record . وإذا كانت هذه الإصدارة معنية بالكتب ، فلا يزال العمل جارياً لتطوير إصدارة أونيكس الخاصة بالمسلسلات والتي تتكون من ثلاث مجموعات من عناصر ما وراء البيانات، خصصت الأولى منها لعنصر العنوان ، والثانية للأعداد الصادرة من المجلة، والثالثة لمعلومات الاشتراكات .

لقد كان الدافع الأساسي لتطوير خطتي أونيكس وإيبيكس هو التزايد المتسارع لمبيعات الكتب الإلكترونية الذي شهده النصف الثاني من تسعينات القرن العشرين الميلادي، والذي حدا بمعظم كبار تجار التجزئة للمواد الإلكترونية إلى استحداث صيغ مفضلة لديهم تقبل بموجبها بيانات الناشرين، إلا أن هذه الصيغ لم تكن موحدة، بل جاءت مختلفة بعضها عن بعض، وبشكل كان يضطر معه الناشر الواحد إلى التعامل بأكثر من صيغة واحدة، وفي الوقت نفسه كان تجار تجزئة المواد الإلكترونية وغيرهم من الموزعين يتلقون بيانات كثيرة في شكل رسائل ذات صيغ غير مألوفة لهم ، ورأوا أن إيجاد صيغة واحدة يتبناها جميع الناشرين سيكون له مردود فوري بالنسبة لهم، كما أدرك الناشرون من جانبهم كذلك ضرورة توفير أكبر قدر من التوحيد الذي يكفل توصيل المعلومات المتصلة بالسوق إلى بائعي الكتب على الإنترنت أو على الخط المباشر online booksellers بسرعة ودقة ، كما

أدركوا كذلك أن بيع التجزئة للمواد الإلكترونية على الوب web e-tailing أدى بالفعل إلى تعويض الافتقار إلى وجود الكتب الفعلية القابلة للتصفح عن طريق توفير بدائلها التي تأتي في شكل صور ومعلومات عن محتواها . وقد أشير في مقدمة الطبعة الأولى من أونيكس إلى أن الكتب التي أتيحت صور لأغلفتها ، فضلاً عن مراجعاتها والمعلومات عنها ، في مواقع متاجر الكتب على الإنترنت online bookstores فاقت مبيعاتها مبيعات الكتب التي لم تتوافر لها مثل تلك المعلومات ، وبمعدل وصل إلى ثمانية كتب لصالح الفئة الأولى في مقابل كتاب واحد للفئة الثانية.

ومن منطلق الأهداف المتعددة التي صممت من أجلها تسجيلة أونيكس، يمكن النظر إلى المعلومات التي تتضمنها أي تسجيلة من تسجيلات منتج أونيكس على أنها مجموعة كبيرة غير مقيدة من المعلومات التي توجد في التسجيلة الببليوجرافية المتداولة في المكتبات التقليدية، وفضلاً عن ما وراء البيانات الوصفية التي تعين القارئ على البحث عن أحد الأعمال والتعرف إلى هويته، تشتمل تسجيلة منتج أونيكس على معلومات أخرى ذات صبغة دعائية تهدف إلى تشجيع القارئ على شراء ذلك العمل، كما تشتمل كذلك على معلومات أخرى ذات صبغة تجارية يفيد منها كل من البائع أو الموزع لهذا العمل.

وعلى اعتبار أن أونيكس يعد أحد معرفات نوع الوثيقة في صيغة لغة الترميز الموسعة XML DTD ، فإنه يعرف كل عنصر من عناصر البيانات عن طريق اسمين يسجل كل منها في تاجين ، يتمثل أحدهما في اسم نصي pneumonic (مثل <PublisherName>)، ويتمثل الآخر في قيمة قصيرة

مرمزة ، ويعرف الأول بالاسم المرجعي reference name ، في حين يعرف الثاني بالتاج القصير short tag . أما الإصدارة رقم ٢,٠ من أونيكس فقد عّرفت كل عنصر من خلال خمس خصائص ، جاءت على النحو التالي : " الوصف غير المقيد"، و" الصيغة "، و"قائمة الرموز " ، و"الاسم المرجعي" ، و"التاج القصير" .

وتتضمن خاصية " الصيغة " كلاً من نوع البيانات، وطولها، مثال ذلك الصيغة التالية : " Fixed length-two numeric digits - أي طول ثابت مكون من رقمين " أو" variable length text, suggested maximum ٢٠٠ characters - أي نص متغير الطول بعدد حروف يقترح أن يكون ٢٠٠ حرفٍ كحد أقصى " . أما خاصية " قائمة الرموز " فتحيل إلى القوائم الاستنادية الخاصة بالقيم المرمزة ، والتي قد تكون قوائم خارجية، أو قوائم محددة في أونيكس نفسه ، مع توضيح كل عنصر عن طريق مثال يبينه . ويعرض الشكل رقم (١٣/ ١) تعريفاً لعنصر "رمز نوع مصدر التسجيلة Record Source Type Code " مأخوذ من " إرشادات معلومات منتج أونيكس _ الإصدارة ٢,٠ .

وتجدر الإشارة إلى أن خطة أونيكس تتميز بتركيزها الشديد على القيم المرمزة ، وعلى استخدام التركيبات (أي مجموعة عناصر بيانات تتواجد مع بعضها البعض) ، وعلى استخدام الخيارات المتعددة لتمثيل البيانات نفسها. وتبين معالجتها للترقيمات الأساسية جميع هذه الخصائص الثلاثة ؛ فهناك عناصر ما وراء بيانات مستقلة معرفة للتعامل مع كل من الرقم المعياري الدولي للكتاب ISBN ، وEAN-١٣ ، وUPC ، هذا فضلاً عن

بعض أرقام المنتجات الأخرى ، ومع ذلك يمكن بدلاً منه استخدام تركيبة مكونة من ثلاثة عناصر، بحيث يأخذ العنصر الأول منها ، وهو رمز نوع معرف المنتج Product Identifier Type Code ، قيمة مرمزة لنوع المعرف الذي يتبعه (مثال ذلك القيمة " ٠٢" التي تعني الترقيم الدولي الموحد للكتاب).

أما العنصر الثاني، وهو اسم نوع المعرف Identifier Type Name ، فهو عبارة عن جملة نصية ، لكنها تستخدم فقط عندما يعّرف رمز النوع خطة خاصة غير معيارية ، أما العنصر الثالث، وهو قيمة المعرف Identifier Value فيكتب فيه الترقيم الأساسي الحقيقي . ويلاحظ في هذا الصدد أن هذه التركيبة يمكن أن تتكرر لتمثيل أي عدد من الترقيمات الأساسية.

إن اقتران رمز نوع المعرف بعنصر قيمة المعرف هو أساس تركيبة أونيكس المثالية، وإن كانت هنالك نماذج أخرى شائعة يتم فيها إقران قيم مرمزة بالعناصر لقيم نصية مكافئة، إلا أن استخدام النموذج الأول يعد إجباريًا، في حين يمكن استخدام النموذج الثاني إما اختيارياً أو في بعض الحالات الخاصة؛

An ONIX code which indicates the type of source which has issued the ONIX record. Field PR 1.5 is optional, and may occur once and only once, independently of the occurrence of any other field.

Format	Fixed-length, two numeric digits
Code list	00 Unspecified (default value) 01 Publisher 02 Publisher's distributor (use only for a distributor appointed by the publisher, as distinct from a wholesaler) 03 Wholesaler 04 Bibliographic agency 05 Library bookseller
Reference name	<RecordSourceType>
Short tag	<a194>
Example	01

الشكل رقم (١/١٣) عينة لتسجيلة أونيكس باستخدام أسماء مرجعية .

المصدر: ONIX Product Information Guidelines, Release ٢,٠<Product Record>

فهنالك على سبيل المثال عنصران في المواصفة مخصصان للقيم المرمزة التي تصف شكل العمل ، يعرف أولهما بشكل المنتج Product Form ، ويعرف الثاني بتفاصيل شكل الكتاب Book Form Detail ، كما أن هنالك عنصراً ثالثاً ، هو عنصر وصف شكل المنتج Product Form Description ، الذي قد يشتمل على وصف نصي لشكل المادة ، إلا أن استخدامه لا يتم إلا عندما تكون القيم المرمزة في العنصرين الآخرين غير كافية لوصف العمل .

كذلك تقدم معالجة بيانات العنوان مثالاً آخر على مجموعة الخيارات التي يتيحها أونيكس ؛ حيث يمكن تمثيل العنوان ببساطة عن طريق استخدام عنصر العنوان المميز للمنتج " Distinctive Title of the Product " هكذا على النحو التالي:

< DistinctiveTitle > The genetics of the dog</ DistinctiveTitle >

ويمكن بدلاً من ذلك استخدام تركيبة تشتمل على عنصرين ، أحدهما

لنوع العنوان المرمز (ويلاحظ في المثال التالي أن القيمة "٠١ " تعني عنواناً مميزاً) ، وثانيهما لنص العنوان نفسه ، هكذا :

<TitleType>٠١</TitleType>

<TitleText> The genetics of the dog <TitleText>

وعندما تكون هناك حاجة للإشارة إلى عدم احتساب بعض الكلمات عند إجراء عملية الصف ، يمكن استخدام شكل آخر ، هكذا :

<TitlePrefix>The</TitlePrefix>

<TitleWithoutPrefix> genetics of the dog </TitleWithoutPrefix>

وفي جميع الحالات، يمكن - على سبيل الاختيار - الإشارة إلى طريقة كتابة الحروف الاستهلالية الكبيرة المستخدمة عن طريق " مؤشر حالة النص Text Case Flag " الذي يحتوي على قيم لحالة الجملة، أو حالة العنوان، وجميع الحروف الكبيرة ، هكذا على النحو التالي :

<TitleType>٠١</TitleType>

<TitleCaseFlage>٠٢</ TitleCaseFlage >(١)

<TitleText> The Genetics of the Dog </TitleText>

ولعل في ذلك ما يبين سمة عامة لخطة أونيكس ، فبصرف النظر عن الاستخدام الإجباري للقيم المرمزة، تتجنب هذه الخطة تحديد قواعد لاختيار

(١) تعني القيمة "٠٢" في المثال أن جميع الكلمات الأساسية في العنوان تبدأ بالحروف الكبيرة [المترجم].

المحتوى أو صياغته ، ولكنها بدلاً من ذلك تسمح بتمثيل البيانات بطرق متعددة، ومن خلال تقديم عناصر تبين بشكل صريح طريقة التمثيل المستخدمة .

وكما أشرنا من قبل ، تتضمن خطة أونيكس العديد من عناصر البيانات المشابهة للعناصر المستخدمة في الفهرسة التقليدية للمكتبات، مثال ذلك الترقيمات الأساسية، وأسماء المؤلفين وغيرهم من المشاركين في العمل، وبيانات الطبعة ، وبيانات النشر، والوصف المادي، والجمهور، والموضوع، إلا أن العديد من هذه العناصر تعالج بشكل مختلف تماماً عما تعالج به في فهرسة المكتبات ؛ فعنصر "التأليف " - على سبيل المثال - يتم تمثيله عن طريق تركيبة "المشارك Contributor" التي يجب أن تشتمل على رمز دور المشارك ، كما أن هناك عناصر محددة لتمثيل الأسماء الشخصية ، إما بالطريقة العادية غير المقننة، أو بالطريقة التي يُقلب فيها الاسم ولكن بشكل غير مقنن كذلك، أو بصيغة مقننة تحدد فيها أجزاء الاسم بشكل مستقل بعضها عن بعض؛ ومن ثم فإن اسم الشخص الواحد يمكن أن يأخذ أكثر من شكل واحد داخل التركيبة نفسها، وهناك كذلك عنصر واحد غير مقنن مخصص لأسماء الهيئات عندما ترد كمشاركة في العمل ، أما عناوين المؤتمرات فلا يمكن إدخالها ضمن فئة المشاركين برغم وجود عناصر لتمثيل عناوين المؤتمرات المرتبطة بالمطبوع، كذلك ليس هناك إرشادات محددة لصياغة أسماء الأشخاص أو الهيئات، كما لم يشر المعيار كذلك إلى أية قوائم استنادية يمكن الاعتماد عليها في تقنين هذه الأسماء .

أما المصطلحات الموضوعية فتقسم في سياق هذه الخطة إلى فئتين : "الموضوعات الرئيسية" ، و"الموضوعات الإضافية" . وبالنسبة للموضوعات الرئيسية لا يمكن أن تستخدم سوى الخطط الموضوعية المعترف بها كمعايير لتجارة الكتب في بلدان أو مناطق معينة ، وهناك لغتان موضوعيتان سائدتان في عالم تجارة الكتب ، الأولى تعرف بقائمة بيسيك BASIC (Book And Serials Industry Communication) المتداولة في الولايات المتحدة بشكل أساسي، والثانية قائمة بيك BIC (Book Industry Communication) المتداولة في المملكة المتحدة ، وفي سياق الطريقة المثالية لأونيكس هناك مجموعات مستقلة لعناصر الموضوع الرئيسي معرفة خصيصاً لمصطلحات هاتين القائمتين . أما بالنسبة لمصطلحات الخطط الأخرى فيمكن استخدام تركيبة الموضوع الرئيسي التي يتم فيها إقران رموز اللغة الموضوعية مع المصطلح الموضوعي نفسه، ويمكن في هذه الحالة إقرار موضوع رئيسي واحد فقط، وتشجيع استخدامه بشكل قوي، مع أن ذلك لا يعد إلزاماً . أما تركيبة الموضوعات الإضافية فيمكن استخدامها لإضافة مصطلحات موضوعية إضافية، ولتسجيل موضوعات من لغات أخرى ، مثل : تصنيف مكتبة الكونجرس ، وقائمة رؤوس موضوعات مكتبة الكونجرس، وتصنيف ديوي الكامل أو المختصر .

أما المعلومات الدعائية فترد أساساً في مجموعتين من العناصر ، الأولى تخص الأوصاف النصية ، والأخرى تخص الربط بالصور والمواد الصوتية

والمرئية . وبالنسبة للأوصاف النصية فيمكن أن ترد في عنصري " التعليق Annotation " أو " الوصف الأساسي Main Description " ، أو في تركيبة " نص آخر Other Text " أو في كلتيهما ، وهذه التركيبة تشتمل على عنصر خاص بالنص نفسه ، وقيمة مرمزة تبين طبيعة النص (مثل قائمة محتويات ، أو مراجعة ... إلخ) ، وقيمة مرمزة أخرى تبين الصيغة التي يرد فيها النص (مثل صيغة آسكي ASCII ، أو صيغة لغة ترميز النص الفائق.. وغيرهما) ، وعنصر خاص بمؤلف النص المصدري ، وعنوان هذا النص ، ثم عناصر لعملية الربط بالأوصاف النصية الخارجية .

وتجدر الإشارة في هذا السياق إلى أن جميع أنواع النصوص يمكن وصفها بدءاً من الإصدارات ذات الأغلفة المطوية للداخل flap cover (١)، والإصدارات ذات الأغلفة الخلفية back cover (٢)، إلى الأوصاف الخاصة بالفئات المختلفة للجمهور (مثل البائعين ، والمطابع، والمدرسين ، والناشرين ، والمكتبات)، ثم إلى النص الكامل لأحد فصول العمل أو للعمل كله . كذلك هناك عناصر أخرى خاصة بالمعلومات الدعائية الصريحة عن المبيعات مثل: وصف الحملات الإعلانية، ومعلومات منتدى تبني الكتاب book club adoption . أما التركيبة الخاصة بعملية الربط فتتضمن العناصر الخاصة بنوع العمل المستهدف (مثل برنامج العروض التقديمية، وصورة الغلاف الأمامي، وغيرهما)، وشكله، ودرجة وضوحه، ونوع الرابط ،

---

(١) حافتا جلدة الكتاب اللتان تثنيان للداخل كما في حالة الكتب المجلدة بالجلد [المترجم] .

(١) يعرف أحياناً بالغلاف السفلي للكتاب، ويعرف بمصطلحين آخرين : reverse cover ، lower cover [المترجم] .

والرابط نفسه، ثم المعلومات الخاصة بمتطلبات التحميل، بما فيها معلومات عناوين الصور captions (١)، والاعتمادات credits (٢)، وحقوق التأليف والنشر ، وشروط الاستخدام .

أما المعلومات التجارية في تسجيلة منتج أونيكس فتضم مجموعات من العناصر التي تقدم معلومات عن وصف الحقوق، والمورِّد الأساسي، والسعر ومدى توافر العمل، بل قد تتضمن معلومات خاصة بسياسية إرجاع العمل، ومواعيد توفيره، والخصومات التي تمنح في حالات الشراء بالجملة batch bonuses ، ونسبة الضرائب الخاصة بالدول الأوربية. ويعرض الشكل رقم (١٣/ ٢) مثالاً لتسجيلة أونيكس باستخدام الأسماء المرجعية، مقتبسة من إرشادات معلومات منتج أونيكس، الإصدارة ٢,٠ .

وتتضمن المواصفات الأخرى لأونيكس كلاً من تسجيلة السلاسل الرئيسية وتسجيلة السلاسل الفرعية (وهما مجموعتان إرشاديتان طورتا لمجموعة مستخدمي أونيكس الألماني German ONIX ولم يطبقا في أي مكان آخر)، وهناك كذلك خطة أونيكس للمسلسلات التي لا تزال في نسختها المسودة، والتي تحدد ثلاث صيغ للتسجيلات : الأولى تسجيلة عنوان العمل المسلسل Serial Title، وتعنى بتقديم معلومات ببليوجرافية عن عناوين المسلسلات بشكل متوافق مع تسجيلة منتج أونيكس، هذا بالإضافة إلى

(٢) لهذا المصطلح عدة عناوين ، منها عنوان النص ، عنوان الصورة ، التعليقات والشروح لصورة في مجلة [المترجم].

(٣) تبصرة في فهرسة الصور المتحركة تدرج الأشخاص (غير هيئة التمثيل) الذين يقومون بمجهود في الإنتاج الفني أو غيره للعمل ، مثل المنتج أو المخرج ...إلخ [المترجم].

عناصر إضافية تتيح معلومات عن كيفية الاتصال بناشر العمل المسلسل، وتاريخ النشر، والسعر، وكيفية الحصول عليه ، ومع أن المواصفات تذكر أنه إذا نشر العمل المسلسل في نسخة مطبوعة وأخرى إلكترونية ينبغي إنشاء تسجيلة مستقلة لكل منهما ، إلا أنها لم تحدد الحالات التي يتحتم فيها إنشاء تسجيلات مستقلة للإصدارات أو الصيغ المختلفة للمطبوع الإلكتروني، ومتى ينبغي أن يتم ذلك ؛ ولعل السبب في ذلك يعود إلى التسليم الذي يقضي بأن أي مطبوع مسلسل يحمل رقما فريداً من الأرقام المعيارية الدولية للمسلسلات ISSN يجب أن تُنشأ له تسجيلة أونيكس مستقلة .

ومن الواضح أن التغيرات التي تصاحب العناوين ستعالج في الإصدارات المستقبلية للمواصفة من خلال توسعة تركيبة " تاريخ النشر"، حيث يفترض أن يتم جمع المعلومات الخاصة بالناشرين السابقين والعناوين السابقة في التسجيلة الخاصة بالعنوان الحالي للعمل المسلسل .


```
<Product>
    <RecordReference>1234567890</RecordReference>
    <NotificationType>03</NotificationType>
    <ISBN>0816018358</ISBN>
    <ProductForm>BB</ProductForm>
    <DistinctiveTitle>British English, A to Zed</DistinctiveTitle>
    <Contributor>
        <ContributorRole>A01</ContributorRole>
        <PersonNameInverted>Schur, Norman W</PersonNameInverted>
        <BiographicalNote>A Harvard graduate in Latin and Italian literature, Norman
        Schur attended the University of Rome and the Sorbonne before returning to the
        United States to study law at Harvard and Columbia Law Schools.  Now retired
        from legal practice, Mr Schur is a fluent speaker and writer of both British and
        American English</BiographicalNote>
    </Contributor>
    <EditionTypeCode>REV</EditionTypeCode>
    <EditionNumber>3</EditionNumber>
    <LanguageOfText>eng</LanguageOfText>
    <NumberOfPages>493</NumberOfPages>
    <BASICMainSubject>REF008000</BASICMainSubject>
    <AudienceCode>01</AudienceCode>
    <ImprintName>Facts on File Publications</ImprintName>
    <PublisherName>Facts on File Inc</PublisherName>
    <PublicationDate>1987</PublicationDate>
    <Height>9.25</Height>
    <Width>6.25</Width>
    <Thickness>1.2</Thickness>
        <MainDescription>BRITISH ENGLISH, A TO ZED is the thoroughly updated,
        revised, and expanded third edition of Norman Schur's highly acclaimed
        transatlantic dictionary for English speakers.  First published as BRITISH SELF-
        TAUGHT and then as ENGLISH ENGLISH, this collection of Briticisms for
        Americans, and Americanisms for the British, is a scholarly yet witty lexicon,
        combining definitions with commentary on the most frequently used and some
        lesser known words and phrases.  Highly readable, it's a snip of a book, and one
        that sorts out – through comments in American – the "Queen's English" –
        confounding as it may seem.</MainDescription>
        <ReviewQuote>Norman Schur is without doubt the outstanding authority on the
        similarities and differences between British and American English.  BRITISH
        ENGLISH, A TO ZED attests not only to his expertise, but also to his undiminished
        powers to inform, amuse and entertain. – Laurence Urdang, Editor, VERBATIM,
        The Language Quarterly, Spring 1988 </ReviewQuote>
    <SupplyDetail>
        <SupplierSAN>1234567</SupplierSAN>
        <AvailabilityCode>IP</AvailabilityCode>
    <Price>
        <PriceTypeCode>01</PriceTypeCode>
        <PriceAmount>35.00</PriceAmount>
    </Price>
    </SupplyDetail>
</Product>
```

الشكل رقم (١٣/ ٢) مثال لتسجيلة أونيكس باستخدام الأسماء المرجعية ، مقتبسة من إرشادات معلومات منتج أونيكس ، الإصدارة ٢,٠ .

أما الصيغة الثانية فهي "تسجيلة أونيكس للوحدة المسلسلة ONIX Serial Item " التي صممت لتوصيل المعلومات المتصلة " بالأجزاء المنفصلة " من المطبوعات المسلسلة ، والتي عادة ما تكون أعدادًا، وأحيانًا تكون مقالات أو وحدات أخرى . وتتمثل الاستخدامات المحتملة لهذه التسجيلة في تقديم خدمات الإحاطة الجارية، وفي إجراءات تسجيل المسلسلات بالمكتبة library check-in. وتشتمل هذه التسجيلة الخاصة بالوحدة المسلسلة على معرِّف للوحدة التي يتم وصفها، بالإضافة إلى معلومات عن عنوان المسلسل المحتوي لهذه الوحدة containing serial number ، وإن كان الكم الأكبر من المعلومات يرد في تركيبتي "العدد Journal Issue "، و" وحدة المحتوى Content Item " ؛ حيث تشتمل تركيبة "العدد" على تركيبات وعناصر تعرف بالعدد الصادر من المجلة من حيث ترقيمه، وتاريخه ، وعنوانه، ونوعه ، وعدد صفحاته ، وفي حالة احتواء العدد على بحوث مقدمة لأحد المؤتمرات ، فيجب في هذه الحالة تقديم معلومات عن هذا المؤتمر . وتجدر الإشارة إلى أن هذه التركيبة - أي تركيبة عدد المجلة - من التركيبات القابلة للتكرار، ويمكن استخدامها في التعريف بالعدد السابق الأخير، وبالعدد الحالي ، وهي مزية لها أهميتها في إجراءات تسجيل الأعمال المسلسلة بالمكتبة . أما تركيبة " وحدة المحتوى " فتستخدم لوصف المواد التي يتضمنها العدد

وصفا تفصيليا، مثل : المقالات ، والافتتاحيات، والمراجعات، كما أن هذه التركيبة تحتوي على تركيبات وعناصر ، مثل عدد الصفحات، والعنوان، والمساهمين، والموضوعات، والتواريخ ذات الصلة، ومعلومات المؤتمرات ذات الصلة .

أما تسجيلة حزمة الاشتراك subscription package record - وهي الصيغة الثالثة - فتعمل على نقل المعلومات عن الاشتراكات المجمعة التي تقدم فيها مطبوعتان مسلسلتان أو أكثر بسعر واحد. وتشتمل هذه التسجيلة على مجموعتين من العناصر تطبق على الحزمة بأكملها، حيث تقدم في سياق المجموعة الأولى المعلومات التي تحدد هوية حزمة الاشتراك من خلال رقم تعريف، وعنوان (غالبًا ما يعدان محليًا لأحد الناشرين أو الوكلاء)، ثم معلومات عن السعر، وإجراءات التوريد . أما بالنسبة للعناوين الفردية التي تقدم بداخل الحزمة فتوصف من خلال عناصر موجودة في تركيبة " حزمة العنوان Title Package " التي تتكرر لكل عنوان من هذه العناوين، حيث تعرف تركيبة حزمة العنوان هذه، العنوان عن طريق رقم أساسي وعنوان . وداخل هذه التركيبة كذلك، تعرف تركيبة " عدد المجلة a Journal Issue " تاريخي البداية والنهاية لمدى تغطية هذا العنوان داخل الحزمة .

أما مواصفة أونيكس للمسلسلات ONIX for Serials - التي صدرت لها مؤخرًا مسودة أحدث من مواصفة تسجيلة منتج أونيكس - فتقدم بعض التوجيهات التي تعمل في ظلها مواصفات أونيكس ، حيث تفضل هذه المواصفة الاعتماد المكثف على التركيبات العامة أكثر من اعتمادها على

تخصيص عناصر محددة لأنواع معينة من القيم، ومن أمثلة ذلك أنه لا يوجد في سياق هذه المواصفة عنصر معرف خصيصًا للرقم المعياري الدولي للمسلسلات، ولا سبيل لتمثيل هذا الرقم سوى تركيبة "معرف عنوان المسلسل Title Identifier Serial " التي تشتمل على قيمة مرمزة لنوع المعرف مقرونة بقيمة المعرف نفسه . كذلك ترى هذه المواصفة استخدام خصائص لغة الترميز الموسعة لأنواع القيم التي كانت قد عرفت على أنها تيجان لغة الترميز الموسعة في مواصفات سابقة ، كما هو الحال مع " مؤشر حالة النص Text Case Flag "، هكذا :

<TitleText textcase="٠٢">Title in Title Case</TitleText>

وإذا كان الناشرون قد طوروا تسجيلة منتج أونيكس في الأساس لتوفير معلومات منتجاتهم إلى بائعي الكتب وغيرهم من الموزعين على الخط المباشر، فسرعان ما رأت المكتبات إمكانية الإفادة مما توفره أونيكس من بيانات، وإذا كان موردو النظم الآلية للمكتبات قد حرصوا كذلك على إضافة إمكانات عرض قوائم المحتويات والمراجعات والأعمال الفنية وغيرها من المحتويات الثرية إلى فهارسهم المتاحة على الخط المباشر، فقد رأوا في تسجيلات أونيكس التي يعدها الناشرون مصدرًا يمكن الاعتماد عليه في توفير هذا المحتوى . كذلك اكتشفت مكتبة الكونجرس أن ثمة عدة استخدامات محتملة لأونيكس، منها الإفادة من معطيات أو مغذيات أونيكس ONIX feeds في توفير البيانات التي تدعم عملية الفهرسة أثناء النشر، وكذلك إجراء عملية الربط بين تسجيلات الفهارس وقوائم المحتويات

المستخلصة من أونيكس[١]، كما رأت كذلك إمكانية ربط المعلومات الببليوجرافية في تسجيلة منتج أونيكس بمارك ٢١، بالرغم من أن التسجيلة [الناتجة عن أونيكس] لا يمكن استخدمها في فهارس المكتبات إلا بعد إجراء قدر كبير من التعديلات عليها، نتيجة لافتقار أونيكس في المقام الأول للضبط الاستنادي للأسماء من ناحية، ولاستخدام أنماط استنادية مختلفة لعدد من عناصر البيانات الأخرى من ناحية أخرى . فعلى سبيل المثال يعتمد أونيكس بالنسبة لقيم عنصر دور المشارك على " قائمة رمز دور المشارك" المعرفة من قبل قائمة "ONIX-defined 'Contributor Role Code List "، بينما يعتمد مارك ٢١ في هذا الصدد على " روابط مارك MARC Relator " . يضاف إلى ذلك أنه يصعب استنباط ما تشير إليه القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية من ضرورة التمييز بين المدخل الرئيسي والمدخل الإضافي من خلال ما يتوافر من معلومات في عناصر أونيكس الخاصة بالمشارك ودوره، كما يصعب كذلك تطبيق إجراءات مارك ٢١ لإنشاء [تقنينات] الحقول الفرعية المناسبة على الجمل النصية الموضوعية غير المتميزة التي يعتمدها أونيكس .

وإذا كان أونيكس قد حظي فور ظهوره بترحيب الناشرين وبائعي الكتب باعتباره معيارًا يلبي حاجات ملحة خاصة بهم ، فقد حظي كذلك بترحيب

(١) LC Cataloging Newsline: Online Newsletter of the Cataloging Directorate Library of Congress ٩, no. ١٢ (November ٢٠٠١), available at

http://www.loc.gov/catdir/lccn/Iccn٠٩١٢.l. Accessed ٣ July ٢٠٠٢.

المكتبات لما رأت فيه مصدراً يوفر لها بيانات أولية وثرية تستخدمها في فهارسها وفي غيرها من النظم الأخرى ، ولهذا فمن المؤكد أن استخدامه سيتزايد بشكل سريع، كما أن مجموعة مواصفاته ستتطور باستمرار، فضلاً عن أنها ستحظى بالرعاية، ومع ذلك، يجب ألا ننسى أن أونيكس لا يزال خطة غير ناضجة ، وليس لها ذلك التاريخ المديد من رصيد الاستخدام في بيئات الإنتاج ، كما يحتمل أن تخضع مواصفاتها للتعديلات كما استخدمت في تبادل كميات أكبر من البيانات في سياق التعاملات التجارية ، كما لا نستبعد أن نرى في المستقبل القريب تطوير بروفايلات أونيكس، ذلك أن كثيراً من مطبقي أونيكس متفقون على ضرورة استخدام مجموعات محددة من خياراته للارتقاء بأنشطة التشغيل البيني .

READINGS :

- Association for Library Collections and Technical Services. Committee on Cataloging: Description and Access. Task Force on ONIX International. Final Report. Available at http://www.ala.org/alcts/organization/ccs/ccda/tf-onix٣.html.
- The CC:DA Task Force on ONIX International issued this report on ONIX in December ٢٠٠١. Although based on ONIX International ١,٢,١, the report is an excel- lent study of ONIX characteristics, implementations, and library application.
- Barnes&Noble.com (home page). Available at http://www.bn.com. Accessed ٢٧ August ٢٠٠٢.

It is interesting to study the way some booksellers have implemented ONIX. Barnes and Noble provides documentation to help publishers contribute their ONIX-formatted data in a way that the bookseller can accept. This documentation is available on the "help desk" portion of the Barnes and Noble website at http://www.barnesandnoble.com/help/pubsub mit-onixl.asp.

- EDItEUR (home page). Available at http://www.editeur.org. Accessed ٢٧ August ٢٠٠٢.

  This website links to guidelines and XML DTDs for the ONIX family of specifications.

  Library of Congress, Network Development and MARC Standards Office. ONIX to MARC٢١ Mapping (December ٢٠٠٠). Available at http://lcweb.loc.gov/marc/onix٢marc.html.

الفصل الثاني عشر

ما وراء البيانات للأغراض التعليمية

شهدت السنوات الأخيرة قدراً كبيراً من النشاط في تطوير خطط ما وراء بيانات متخصصة لوصف المواد التعليمية. ويلقي هذا الفصل الضوء على حقول مارك٢١ الخاصة بمعلومات المناهج الدراسية ، ومجموعة عناصر ما وراء بيانات بوابة المواد التعليمية الخاصة بخطط الدروس وما شابهها من مصادر المناهج الدراسية ، وخطة جمعية المهندسين الكهربائيين و مهندسي الإلكترونيات لما وراء بيانات الكيان التعليمي الخاصة بمصادر التعلم.

القواعد الأنجلو أمريكية للفهرسة ومارك **AACR٢/MARC** :

تعد الكتب وأشرطة الفيديو والبرمجيات وغيرها من المواد التي تقتنيها المكتبات من أكثر المواد استخداماً في الأغراض التعليمية ، وممن الممكن استخدام الفهرسة المقننة اعتمادًا على القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية وصيغة مارك لدمج هذه المواد في فهرس المكتبة. وفي عام ١٩٩٣م حدثت عدة تغيرات على مارك الأمريكي USMARC بناء على طلب من مؤسسة شمالي غرب أوهايو لتقنيات التعليم وشركائها للارتقاء بإجراءات تسجيل المعلومات المتعلقة بالمناهج الدراسية، ويتضمن مارك عدة حقول ذات علاقة مباشرة بالجوانب التعليمية، منها : تبصرة الملخص Summary (الحقل رقم ٥٢٠)، وتبصرة الجمهور المستهدف Target Audience (الحقل ٥٢١)، وتبصرة معلومات برنامج الدراسة Study Program Information (الحقل ٥٢٦) ، وحقل المصطلح الكشفي - هدف المنهج الدراسي Index Term-Curriculum Objective (الحقل ٦٥٨). وفي بعض الأحيان تسمى تسجيلات مارك التي تشتمل على هذه الحقول سي مارك CEMARC اختصاراً " لمارك الموائم للمناهج

الدراسية curriculum-enhanced MARC "، ولهذه الحقول أهميتها الخاصة لمعلمي الفصول الدراسية الأولية K-١٢ .

ويستخدم الحقل ٥٢٠ (الخاص بتبصرة الملخص ، والمستخلص، والشروح ، والمجال ... إلخ) مع المؤشر الأول "١ " لتسجيل النص الكامل لمراجعة ما أو لعرض معين أعد لأحد الكتب أو البرامج أو أي مصدر آخر من المصادر التي تخضع للوصف . وبرغم ندرة تضمين موظفي مراكز مصادر التعلم لتلك المراجعات في تسجيلات الفهرسة؛ فإنها غالباً ما ترد متضمنة في التسجيلات التي يعدها موردو مثل هذه المصادر.

أما الحقل ٥٢١ (تبصرة الجمهور المستهدف) فيمكن استخدامه لتسجيل معلومات عامة عن الجمهور المستهدف ، ومن أمثلة هذه التبصرات ، تبصرة تقول على سبيل المثال : " برنامج مصمم للجغرافيين والمخططين والجيولوجيين وعلماء الأرصاد الجوية وغيرهم ممن لديهم اهتمام مهني بتحليل البيانات المكانية" (يلاحظ أن هذا المثال وجميع الأمثلة الأخرى في هذا الفصل مأخوذة من صيغة مارك٢١ المختصرة للبيانات الببليوجرافية MARC٢١ Concise Format for Bibliographic Data). كما يمكن استخدام هذه التبصرة كذلك للمعلومات الخاصة بالمستوى القرائي، والمستوى العمري ، والمستوى الدراسي ، والسمات الخاصة بالجمهور، ومستوى الدافعية. مع ملاحظة أن الحقل الفرعي "a " من هذا الحقل يستخدم لتسجيل الجمهور المستهدف ، بينما يستخدم الحقل الفرعي " b" للإشارة إلى اسم الجهة التي حددت هذا الجمهور المستهدف. وعليه فإن التبصرة التالية المسجلة هكذا: ٥٢١ ٢#$a٧ & up. ، يتبين منها أن مستوى الاهتمام العمري (المحدد في المؤشر الأول بالقيمة "

٢ ") هو السن من سبع سنوات فأكثر.

ويستخدم الحقل ٥٢٦ (الخاص بتبصرة معلومات برنامج الدراسة) لبيان ما إذا كانت المادة تصنف على أنها من إحدى مكونات برنامج دراسي معين ؛ فبرنامج القارئ السريع accelerated reader program من البرامج المشهورة في المدارس الابتدائية والمتوسطة، وتحتسب فيه النقاط على حسب طول الكتب ومستوى صعوبتها، ويقوم التلاميذ باختيار وقراءة الكتب في حدود نقاط معينة، ثم يخضعون لاختبارات عن طريق الحاسب لمدى استيعابهم لما قرأوا. وهنا يمكن استخدام الحقل ٥٢٦ لبيان ما إذا كانت أحد الأعمال جزءاً من برنامج القارئ السريع ، هكذا على النحو التالي:

٥٢٦ ٠#$aAccelerated Reader/Advanced Learning Systems

$b٥,٠$CA.٠ $d٧٥

ويلاحظ في هذا المثال أن الحقل الفرعي " a" يسجل به اسم البرنامج ، في حين يسجل بالحقل الفرعي "b " مستوى الاهتمام ، والحقل الفرعي " c " ، المستوى القرائي ، أما الحقل الفرعي " d" فيسجل به النقاط التي حققها القارئ.

ويستخدم الحقل الفرعي ٦٥٨ (المصطلح الكشفي - هدف المنهج) لتسجيل طبيعة الأهداف التي يسعى المنهج إلى تحقيقها ، بمعنى هل هي أهداف محلية تتصل بالولاية أم أهداف وطنية أم أهداف خاصة. ويكتسب ذلك أهمية كبيرة اليوم بعد أن بات الاختبار المقنن standardized testing المقيد بمواصفات الإنجاز الرسمية هو القاعدة أو المعيار الذي يحتكم إليه في هذا الصدد ، إلا أن ذلك ينطوي إلى حد ما على إشكالية التعبير عن ذلك

بمصطلحات عملية ، بمعنى أن المادة نفسها يمكن أن تلبي العشرات من أهداف المناهج التي تدرس في ولايات مختلفة. وفي سياق هذا الحقل يخصص الحقل الفرعي " a" للهدف الرئيسي ، في حين يخصص الحقل الفرعي " b" للهدف الثانوي، والحقل الفرعي " c" لإحدى التمثيلات المرمزة للهدف ، والحقل الفرعي "d " لمدى قوة الارتباط بالهدف. أما الحقل الفرعي "٢ " فيخصص لتسجيل معيار المنهج الذي يظهر فيه الهدف. وتضم قائمة رموز مارك MARC Code List الخاصة بالروابط Relators والمصادر وتقنيات الوصف قيماً مرمزة لعدد محدود من معايير المناهج. وفي المثال التالي نجد أن الكود "ohco " هو بمثابة اختصار لأهداف المناهج في ولاية أوهايو.

٦٥٨##$aReading objective ١ (fictional)$bunderstanding language, elements of plots,themes,motives,characters,setting by responding to the multiple-meaning word$cNRPO٢-١٩٩١$dhighly correlated.$٢ohco

ومن الملاحظ أن قيم بيانات حقول سي مارك هذه أو مارك المتوافق مع المناهج الدراسية لا تحكمها القواعد الأنجلو أمريكية للفهرسة في طبعتها الثانية؛ إلا أن مارك ٢١ وقائمة رموز مارك تضمنا وصفاً كاملاً لمحتوى هذه القيم ، ولأن الغالبية العظمى من مراكز مصادر التعلم تلجأ إلى شراء تسجيلات الفهرسة، فإن استخدام هذه الحقول يعتمد بشكل كبير على مدى تطبيق الموردين لهذه القيم ، ومن أمثلة ذلك أن مؤسسة فوليت the Follett Corporation تتكفل بإعطاء معلومات عن المستوى القرائي،ومستوى الاهتمام ، ومصادر المراجعة، كما أنها تقوم بوضع المعلومات عن

برنامجي "القارئ النهم أو السريع accelerated reader " و" نقاط القراءة reading counts " في الحقل ٥٢٦.

بوابة المواد التعليمية أو جيم **GEM** :

بالرغم مما جرت عليه العادة من عدم تضمين فهارس المكتبات العديد من المصادر التعليمية ، كمخططات الدروس ومخططات الوحدات والأنشطة الدراسية ، فإن هنالك اهتمامًا كبيرًا بتقاسم هذه المواد على المستويين الإقليمي والوطني، ولهذا السبب طورت عدة مبادرات حكومية وصناعية مشروعات كبيرة لفهرسة هذه المواد التعليمية . وتعد بوابة المواد التعليمية Gateway to Educational Materials التي تبنتها كل من وزارة التعليم الأمريكية ومركز إريك للمعلومات والتقنية في جامعة سيراكوس من إحدى هذه المشروعات ، وتهدف هذه البوابة إلى " توفير سبل الوصول أو الإتاحة الشاملة الجامعة one-stop, any-stop access لجميع المجموعات المعتمدة على الإنترنت من المواد التعليمية الأساسية غير المفهرسة، المتاحة عبر المواقع الفيدرالية الحكومية، والجامعية، والربحية ، وغير الربحية"(١).

وتحقيقًا لهذا الهدف طورت البوابة موقعًا مركزيًا لبحث المصادر التعليمية، كما طورت مجموعة عناصر ما وراء بيانات (الآن في إصدارتها رقم ٢,٠ GEM) مع أدوات برمجية وتوثيقية وتدريبية لمساعدة المشاركين على إنشاء ما وراء البيانات سواء لكشافاتهم المحلية أو للمستودع المركزي

(١) GEM Project Site (home page), available at http://www.geminfo.org/ Accessed ١ July ٢٠٠٢.

للبوابة، على أن تضمن ما وراء البيانات هذه إما في المصادر نفسها عن طريق تيجان الميتا <meta> ، أو بكتابتها في ملفات مستقلة بصيغة لغة ترميز النص الفائق ، بحيث يقوم برنامج الجامع الخاصة بالبوابة GEM harvester في كلتا الحالتين بجمعها وتكشيفها .

وقد عملت مبادرة جيم عن كثب مع مبادرة دبلن كور لما وراء البيانات من أجل تطوير مجموعة العناصر الأساسية الخاصة بها، والمعروفة اختصارًا بجيم، وتتألف هذه الإصدارة - GEM ٢,٠ - من العناصر الخمسة عشر الخاصة بدلن كور مع مقيداتها ، بالإضافة إلى مجموعة عناصر أخرى خاصة بجيم مع مقيداتها كذلك ، وهي ما سنعرف بها في الفقرات التالية .

أول هذه العناصر هو عنصر الجمهور Audience ، ويعرف على أنه " إحدى الفئات التي يوجه إليها المصدر "، ويقابل هذا العنصر إلى حد ما العنصر الموسوم " التبصرة " التي يخصص لها الحقل رقم ٥٢١ في مارك. ولهذا العنصر مقيدان: المستفيد beneficiary، والوسيط mediator ، ويستخدمان للتمييز بين المتعلم المستهدف من ناحية، والمعلم أو المدرب أو أي كيان آخر يقوم بدور الوسيط لإتاحة الوصول للمصدر من ناحية أخرى . كما يلحق بهذا العنصر مقيدات أخرى، هي : المستوى Level - للإشارة إلى المستوى الدراسي الذي تناسبه المادة أو غيره من المستويات الأخرى ، والعمر Age - للإشارة إلى عمر الجمهور المستهدف أو المدى العمري للجمهور الذي يوجه إليه المصدر، ثم عنصر المتطلبات السابقة للتعامل مع هذه المادة التعليميةPrerequisites .

وتجدر الإشارة في هذا الصدد إلى أن عنصر " الجمهور " والمقيد " الوسيط " الخاص به ، سبق أن طرحتهما لجنة العمل الخاصة بالتعليم التابعة لدبلن كور أمام لجنة الاستخدام التابعة لمبادرة دبلن كور لما وراء البيانات DCMI Usage Board ، والتي اعتمدتهما كأول مصطلحين من مصطلحات دبلن كور المتخصصة في مجالات معينة domain - specific Dublin Core terms ، والتي يقصد بها تلك المصطلحات التي لم تثبت حاجة جميع المجالات لها ، وإنما تقتصر الحاجة إليها على مجال محدد دون غيره من المجالات .

ولا تزال مبادرة دبلن كور لما وراء البيانات في حاجة لتوضيح كيف أن المصطلحات المتخصصة بمجال معين والمعتمدة من قبل هذه المبادرة DCMI-approved domain-specific terms تختلف عن الامتدادات الأخرى الخاصة بمجال معين domain-specific extensions والمستخدمة في الخطط القائمة على دبلن كور ، وكذلك لتوضيح الكيفية التي تمثل بها تلك المصطلحات في كل من مكاتب التسجيلregistries ، وتركيبات التبادل .

وهناك عناصر أخرى خاصة بجيم ، هي : عنصر" الفهرسة "، و "الفترة الزمنية" ، و"المصادر الضرورية " ، و"الجوانب التربوية pedagogy " ، و"المعيار" ، مع ملاحظة أن عنصر الفهرسة يختص بتسجيلة ما وراء البيانات نفسها وليس بالمصدر الموصوف ، حيث يسجل به " المعلومات عن الأفراد أو الهيئات أو كليهما ممن أنشأ تسجيلة جيم " ، ولهذا العنصر مقيد يعرف بمقيد " الدور role " يستخدم لتوضيح دور ذلك الكيان في الفهرسة.

أما عنصر "الفترة الزمنية" فيصف الوقت أو عدد الجلسات المطلوب قضاؤها لتحقيق الإفادة من المصدر، كما يخصص عنصر "المصادر الضرورية " ليسجل به قائمة بالمصادر الأخرى المطلوب توفيرها لتحقيق الإفادة من المصدر، مثل معرفة القراءة أو الخلفية الثقافية، وتوفير بعض المستلزمات الفنية (كالمساطر، وأقلام الرسم الملونة، وورق الأشكال التركيبية) .

أما عنصر "الجوانب التربوية" فيخصص للتعريف بالأساليب والإجراءات التربوية ، وله عدة مقيدات هي : التجميع grouping ، وطريقة التدريس TeachingMethod ، والتقييم Assessment ، مع ملاحظة أن قيم جميع هذه المقيدات يجب أن يعتمد في اختيارها على لغة مقيدة معتمدة من قبل جيم ، ويبين المقيد " التجميع " السياق الذي يجب أن تقدم فيه التعليمات ، بدءًا من المجموعة الفردية ، فالمجموعة الصغيرة ، ثم المجموعة الكبيرة المتجانسة ، أما بالنسبة للمقيد " طريقة التدريس " ، فتشمل اللغة المقيدة المعتمدة من قبل جيم على عدة قيم ، مثل : الإجراءات المعملية Lab procedures ، والعرض Demonstration ، والتعلم التطبيقي hands-on learning ، أما مقيد " التقييم " فيسجل الطريقة المفضلة لتقييم الطالب ، مثل : الاختبار المقنن Standardized testing ، والاختبار العادي Testing ، والتقييم الذاتي Self-evaluation.

ويخصص عنصر "المعيار" ـ المشابهة للحقل ٦٥٨ في مارك - لربط المصدر الموصوف بالمعايير التعليمية على المستوى المحلي أو الوطني أو بكليهما، غير أن هذا العنصر - وعلى عكس ما عليه الحال بالنسبة للحقل

المقابل له في مارك ، أي الحقل ٦٥٨ ـ لا يقصد به أن يكون مصطلحًا كشفيًا ، وإنما يمكن كتابته كنص حر غير مقنن ، أو كقيمة تركيبية تتضمن كلاً من اسم معيار المنهج الدراسي ، ودرجة الصفة الرسمية للمعيار ، والأهداف الرئيسية والفرعية التي تسعى لتحقيقها . أما المقيد الفردي " الرابط Correlator " فيحدد في سياق اسم الفرد أو الجهة التي ربطت المصدر بالمعيار .

كما يعمل بروفايل جيم على توسعة عناصر دبلن كور عن طريق مقيدات توضيح العناصر ؛ حيث يضيف هذا البروفايل مقيد " الدورRole " لكل من عنصر المنشئ ، والمشارك ، والناشر ، كما يضيف المقيدين " تاريخ الإتاحة على الخط المباشرPlacedOnline " ، وتاريخ التحديث RecordCreated " لعنصر التاريخ ، والمقيد " رمز السعر Pricecode " إلى عنصر "الحقوق"، والمقيد "منصة أو بيئة التحسيب computing platform " إلى عنصر الشكل والصيغة .

وفضلاً عن ذلك ، يضيف هذا البروفايل ثلاثة مقيدات إلى عنصر " المعرف Identifier " للتمييز بين ثلاثة معرفات هي : المعرفات المعيارية العامة PublicID ، والمعرف الفريد لنظام جيم SID ، والمعرف الفريد للنظام المحلي الذي تستخدمه الجهة المشاركة SDN. كما يضيف أكثر من اثني عشر مقيداً لعنصر " العلاقة relation " للتعريف بالمواد الخارجية ذات الصلة بالمصدر ، مثل : تقديرات المحتوى ، والأشكال المختلفة للمراجعات التي أعدت للمصدر ، ومعدلات الجودة، ومعلومات الطلب ، مع ملاحظة أن أحد هذه المقيدات ، وهو المقيد الموسوم " يتوافق مع conformsTo " ، قد قبله

مجلس الاستخدام التابع لمبادرة دبلن كور لما وراء البيانات ، كمقيد توضيح للعنصر الخاص بمجال محدد ، وقد قصد من ورائه أن يستخدم "للإشارة إلى أحد المعايير المعتمدة الذي يتوافق معه المصدر "

وحددت مبادرة جيم ثلاثة مستويات للفهرسة جاءت على النحو التالي : مستوى بروفايل جيم GEM profile ، والمستوى الأول ، والمستوى الثاني، ويعد مستوى بروفايل جيم الحد الأدنى المقبول للفهرسة ، وهذا المستوى يتألف من العناصر الإجبارية التالية : الفهرسة ، والشكل أو الصيغة ، والجمهور ، والمستوى ، والوكيل المباشر (أي القيمة التي تعطي لدور الناشر) ، والنوع ، والعنوان ، وموضوع جيم، وتاريخ إنشاء التسجيلة ، والحقوق ، والوصف . أما المستوى الأول[1] - وهو المستوى المفضل - فيتألف من العناصر التي يتضمنها مستوى بروفايل جيم بالإضافة إلى عنصري : الكلمات الأساسية، والجمهور.

أما المستوى الثاني[2] فيضم أية عناصر أخرى إضافية تتجاوز حدود المستوى الأول بعناصره . ويعرض الشكل رقم (١٢ / ١) مثالاً للفهرسة في المستوى الثاني مأخوذ من الدليل الإرشادي التدريبي لجيم ، مع ملاحظة أن المثال المعروض في

(١) يعد هذا المستوى هو المستوى الثاني من حيث الترتيب على اعتبار أن مستوى بروفايل جيم هو المستوى الأول [المترجم].

(٢) يعد هذا المستوى هو المستوى الثالث من حيث الترتيب على اعتبار أن مستوى بروفايل جيم هو المستوى الأول [المترجم].

Level Two (Documentation) Cataloging Example

Title:	Your own Classroom Court
GEM Subjects:	Level One: Social Studies Level Two: United States Government
Keyword:	Constitutional law, Courts, Equal [illegible], Trial by jury
Description:	To allow students the opportunity to further their knowledge of the law and its legal proceedings. To experience "trial by a jury of your peers" in simple matters. To give each student a job in the courtroom and to vary these positions throughout the year.
Creator:	Rita Irene [illegible]
Resource Type:	Lesson Plan
Grade Levels:	[illegible]
Audience:	A tool for whom? Teachers Who is the ultimate beneficiary? Students
Pedagogy:	Grouping: Large group instruction Teaching Methods: Hands-on learning, Role Playing
Identifier:	SID: [illegible] SDN: AELP-GOV[illegible]
Date:	Generated by GEMCat
Format:	Text/HTML
Cataloging Agency:	Application: GEMCat (This information is generated by GEMCat) Version: [illegible] (This information is generated by GEMCat) Name: GEM Email: geminfo@geminfo.org URL: http://www.geminfo.org
Online Provider:	Organization/Person Name: AskERIC Email: askeric@askeric.org URL: http://www.askeric.org
Language:	English

التسجيلة والمسمى" محكمة فصلك الدراسي(١) your own classroom court "

(١) حصة دراسية يتم فيها محاكاة لجلسات إحدى المحاكم تهدف إلى تعليم الطلاب إجراءات التقاضي ومراحله وأساليب الدفاع وآلياته [المترجم].

هو رابط مباشر للمصدر نفسه .

وتجدر الإشارة إلى أن جيم لا يتقيد بمجموعة معينة من قواعد المحتوى، غير أن موقعه على الوب يتيح عددًا من الأدوات المساعدة على إنشاء ما وراء بيانات جيم تبنت تطويرها الجهات المشتركة في هذا المشروع ، من هذه الأدوات أحد الأنظمة الفرعية module المجانية للفهرسة الموجود في صيغة لغة الجافا، وفي بعض الإصدارات المتاحة على الوب ، ومنها كذلك الدليل الإرشادي للتدريب، بالإضافة إلى نماذج جاهزة template لإدخال البيانات (تسمى نماذج إخراج أو عرض الصفحات style sheets)، كما تبنى مشروع جيم تطوير قوائم استناد لعدة عناصر، تتضمن الغالبية العظمى من العناصر الخاصة بالجوانب التعليمية وعنصر الموضوع. مع ملاحظة أنه حيثما تتوافر القائمة الاستنادية يكون استخدامها إجباريًا، ومع ذلك يتيح جيم للمشروعات المحلية إمكانية تطوير اللغات المقيدة الخاصة بها، كذلك هنالك إرشادات تقدم الحد الأدنى من التوجيه على كيفية تمثيل قيم البيانات، حيث يرشد نموذج "AskERIC" لإخراج الصفحات - مثلاً - المستفيد إلى القيام بإدخال أسماء المؤلفين بالشكل الذي تظهر به في الوثيقة، مع أن هذه المسألة لم يتعرض لها أي من الدليل الإرشادي التدريبي أو نماذج الإخراج الأخرى .

وتعد مجموعة عناصر جيم مستقلة التركيب من الناحية النظرية، كما يمكن تمثيل جوانبها الدلالية - كما هو الحال بالنسبة للجوانب الدلالية الخاصة بدبلن كور - في عدة صيغ، أما تطبيقيًا فإن التسجيلات التي سيتم

تضمينها في بوابة جيم ، يجب أن تنشأ في صيغة يقبلها مشروع جيم، وإذا ما تم تحويل إحدى قواعد البيانات المحلية الموجودة بالفعل إلى جيم ، فإن دفعة المخرجات يمكن أن تكون في صيغة لغة الترميز الموسعة ، أو في تركيبة جيم معرفة داخليًا تسمى "التركيبة - ١ " syntax-١ ، أو في صيغة لغة ترميز النص الفائق . وإذا ما استخدم النظام الفرعي الخاص بفهرسة جيم عند إنشاء التسجيلات الفردية، فإن المخرجات ستأتي في صيغة لغة ترميز النص الفائق . ويعرض الشكل رقم (١٢ / ٢) مثالاً لتسجيلة في صيغة لغة ترميز النص الفائق أنشئت عن طريق أحد هذه النظم الفرعية للفهرسة .

Level Two (Documentation) Cataloging Metadata

```
<meta name="DC.package.begin" content="1">
<meta name="DC.title" content="(lang=en)Your own Classroom Court">
<meta name="DC.format" content="(scheme=IMT)(type=contentType)text/HTML">
<meta name="GroupStart" content="1">
<meta name="DC.publisher" content="(type=role)onlineProvider">
<meta name="DC.publisher" content="(type=name)AskERIC">
<meta name="DC.publisher" content="(type=email)askeric@askeric.org">
<meta name="DC.publisher" content="(type=homePage)http://www.askeric.org">
<meta name="GroupEnd" content="1">
<meta name="DC.type" content="(scheme=GEM)Lesson plan">
<meta name="GroupStart" content="2">
<meta name="DC.subject" content="(scheme=GEM)(type=levelOne)Social studies">
<meta name="DC.subject" content="(scheme=GEM)(type=levelTwo)United States government">
<meta name="GroupEnd" content="2">
<meta name="GroupStart" content="3">
<meta name="DC.creator" content="(type=namePersonal)Rita Irene Esparza">
<meta name="GroupEnd" content="3">
<meta name="DC.date" content="(scheme=ISO8601:1988)(type=recordCreated)1999-04-01T13:48:53-5:00">
<meta name="DC.description" content="The goal of this lesson is to allow students the opportunity to further their knowledge of the law and its legal proceedings. To experience trial by a jury of your peers in simple matters. To give each student a job in the courtroom and to vary these positions throughout the year.">
<meta name="GroupStart" content="4">
<meta name="DC.identifier" content="(scheme=GEM)(type=SID)AskERIC">
<meta name="DC.identifier" content="(type=SDN)AELP GOV0053">
<meta name="GroupEnd" content="4">
<meta name="DC.subject" content="(type=keywords)Constitutional law,Courts ,Equal protection,Trial by jury">
<meta name="DC.language" content="(scheme=Z39.53)(type=text)English,">
<meta name="GroupStart" content="5">
<meta name="DC.rights" content="(type=priceCode)HF">
<meta name="GroupEnd" content="5">
<meta name="DC.package.end" content="1">
<meta name="GEM.package.begin" content="1">
<meta name="GroupStart" content="6">
<meta name="GEM.cataloging" content="(type=application)GEMCat">
<meta name="GEM.cataloging" content="(type=version)3.21">
<meta name="GEM.cataloging" content="(type=name)GEM">
<meta name="GEM.cataloging" content="(type=email)geminfo@geminfo.org">
```

الشكل رقم (٢/١٢)

تسجيلة ما وراء بيانات جيم في صيغة لغة ترميز النص الفائق أنشأها النظام الفرعي للفهرسة في جيم

المصدر : The GEMCat Training Manual.

ما وراء بيانات الكيان التعليمي IEEE/LOM ، والنموذج المرجعي لكيان المحتوى القابل للمشاركة ADL/SCORM.

في الوقت التي تعنى فيه بوابة المواد التعليمية أو جيم بخطط الدروس ووحدات المناهج الدراسية وما شابه ذلك من مواد تعليمية ، تركز بعض المبادرات الأخرى اهتمامها على نظم إدارة ووصف كيانات التعليم . ومن هذه المبادرة جهود كل من لجنة معايير تقنيات التعليم Learning Teaching Standards Committee (LTSC) التابعة لجمعية المهندسين الكهربائيين ومهندسي الإلكترونيات Electrical and Electronics Engineers (IEEE) ، ونظم الإدارة التعليمية لائتلاف التعليم الكوني أو الشامل IMS Global Learning Consortium ، ومشروع أريادن البريطاني the British ARIADNE (اتحاد شبكات التأليف والنشر التعليمي عن بعد لأوربا Alliance of remote Instructional Authoring and Distribution Networks for Europe) ، ومبادرة التعليم الموزع المتقدم لوزارة الدفاع الأمريكية the U.S

. Department of Defense's Advanced Distributed Learning(ADL)

ومع أن كيانات التعليم عرّفت تعريفات متنوعة ، فإنه ينظر إليها على أنها وحدات صغيرة لمحتوى تعليمي يمكن إعادة استخدامها في سياقات مختلفة . وفي الوقت الذي ترى فيه لجنة معايير تقنيات التعليم التابعة لجمعية المهندسين الكهربائيين ومهندسي الإلكترونيات أن الكيانات التعليمية هي التي تشتمل على المحتوى الرقمي والمحتوى غير الرقمي ، تقتصر معظم التعريفات الأخرى للكيانات التعليمية على الكيانات التي تشتمل على المحتوى الرقمي دون غيره . وعلى أية حال ، فمن المتفق عليه بشكل عام أن الكيانات التعليمية ينبغي أن تكون وحدات ذاتية الاحتواء self-contained units ومؤهلة لأن يعاد استخدامها في أغراض مختلفة ، ومؤهلة كذلك لأن تدمج مع كيانات تعليمية أخرى لينتج بذلك تجميعات تعليمية جديدة.

وقد ركزت تلك المبادرات التي أشرنا إليها في الفقرة السابقة على تطوير خطط ما وراء بيانات فعالة لكيانات التعليم (حتى أن أحد واضعي نظريات التعليم يشكو من إنفاق الكثير من الموارد على تطوير معايير ما وراء البيانات أكثر من إنفاقها على تطوير نظريات تعلم، ويتهكم قائلاً: وستجدنا أمام مكتبات رقمية مليئة بكيانات تعلم يسهل العثور عليها، دون أن نعرف كيف نستفيد منها أو نستخدمها)(١)". ولا شك أن هذا النشاط كان ضروريًا لسببين على الأقل، أولهما أن كيانات التعلم لا تقتصر على الوسيط النصي أو على المادة التي كانت في أصلها مادة نصية فقط، وإنما يمكن أن تكون

---

(١) David A. Wiley, untitled (web document), available at http://wiley.ed.usu.edu/docs/encyc.pdf. Accessed ١ July ٢٠٠٢.

محملة على أي وسيط (رقمي) أو أحد الوسائط المتعددة، ومن ثم تكتسب ما وراء البيانات الخارجية أهميتها في استكشاف هذه الكيانات، وثانيهما أن كيانات التعلم، وباعتبارها أصغر وحدات المحتوى ، تتطلب أن توصف في مستوى من أدق التفاصيل التي لا يمكن أن تحققه الفهرسة التقليدية للمكتبات . ويمكن القول إن خطط ما وراء البيانات القابلة للتطبيق لا ينبغي لها أن تصف هذه الوحدات الصغيرة فحسب ، وإنما عليها أن تؤدي ذلك بطريقة يمكن أن تدفع بالتكامل مع الوحدات الأخرى إلى تحقيق برامج تعليمية هادفة .

كما يمكن القول إن الخطط التي طرحتها هذه المبادرات الثلاث كانت مرتبطة تاريخيًا بعضها ببعض، بل يمكن اعتبارها أشكالاً مختلفة أو بروفيلات لبعضها بعضًا ، فمسودة معيار ما وراء بيانات الكيان التعليمي LOM لجمعية المهندسين الكهربائيين ومهندسي الإلكترونيات اعتمدت أساسًا على مواصفات مبكرة من مشروع أريادن ، ثم طورت بعد ذلك بمعطيات هامة من مبادرة نظم الإدارة التعليمية ، كذلك قام مشروع أريادن بعد فترة من إطلاقه بتعديل مواصفته ليصبح وكأنه بروفايل متوافق مع معيار ما وراء بيانات الكيان التعليمي، كذلك تتكون مواصفة ما وراء بيانات مصدر التعلم لنظم الإدارة التعليمية من [عناصر] ما وراء بيانات الكيان التعليمي ، بالإضافة إلى التعديلات التي أجريت على تلك المواصفة ، والتي أمكن دمجها بعد ذلك في مسودة معيار ما وراء بيانات الكيان التعليمي.

وللنموذج المرجعي لكيان المحتوى القابل للمشاركة لمبادرة التعلم الموزع المتقدم ثلاثة مجموعات من عناصر ما وراء البيانات ، تخص أولاها الوسائط الخام ، وتختص ثانيتها المحتوى ، في حين تختص ثالثتها بالبرامج الدراسية ، ولكل منها بروفايل ما وراء بيانات مصدر التعلم لنظم

الإدارة التعليمية .

ولعل مما تجدر الإشارة إليه في هذا السياق أنه في عام ٢٠٠١م قامت كل من لجنة معايير تقنيات التعلم التابعة لجمعية المهندسين الكهربائيين ومهندسي الإلكترونيات ، ومبادرة دبلن كور لما وراء البيانات بتوقيع مذكرة تفاهم تعبر عن التزامهما المشترك بالتعاون على تطوير ما وراء بيانات قابلة للتشغيل البيني لمجالات التعلم والتعليم والتدريب . وسنركز في الجزء المتبقي من هذا الفصل على التعريف بمسودة معيار ما وراء بيانات الكيان التعليمي لجمعية المهندسين الكهربائيين ومهندسي الإلكترونيات، مع عدم تجاهلنا الحديث عن بعض الخطط الأخرى .

تتكون ما وراء بيانات الكيان التعليمي من عشرات من عناصر ما وراء البيانات ، جمعت في تسع فئات أساسيات هي: " السمات العامة General " - وتضم العناصر التي تعني بوصف الكيان التعليمي بشكل عام ، مثل : العنوان ، والوصف ، والكلمات الأساسية أو المفتاحية . ثم فئة " دورة الحياة Lifecycle " لوصف تاريخ الكيان التعليمي وحالته الراهنة والكيانات الأخرى التي تأثر بها في طور نشأته ، وفضلاً عن المعلومات التي تتصل بنشأة الكيان وحالته ، تشتمل هذه الفئة على عناصر تبين أسماء المشاركين في محتوى العمل ودور كل منهم ، بما في ذلك منشئ العمل ، ومن الملاحظ أن هذا التقسيم يتناقض مع الممارسات الببليوجرافية للمكتبات التي تعامل عملية التأليف على أنها خاصية عامة ، كما أن ذلك ينقل عملية الوصف إلى مرحلة الوصف القائمة على الحدث event-based characterization - كما

وصفها كارل لاجوز Carl Lagoze [(١)]، ثم تأتي بعد ذلك فئة " وصف ما وراء البيانات Meta-Metadata " وتختص بالمعلومات التي تصف عملية فهرسة الكيان التعليمي ، وليس وصف الكيان التعليمي نفسه ، ثم فئة " الخصائص الفنية " للسمات الفنية للكيان التعليمي ، ثم فئة الخصائص التعليمية " لوصف السمات التعليمية أو التربوية للكيان ، ثم فئة " الحقوق " لبيان حقوق الملكية الفكرية وضوابط الاستخدام ، وفئة " العلاقة " لبيان علاقات الكيان بكيانات التعلم الأخرى، ثم فئة " التعليقات " لتوثيق الشروح أو التعليقات التي تتصل بالعمل، وأخيرًا فئة " التصنيف " لتصنيف الكيان ضمن نظم التصنيف التربوية مثل المستوى القرائي .

وتجدر الإشارة إلى أن كل عنصر من عناصر ما وراء البيانات يتم وصفه من خلال سبع خصائص ، هي : "الاسم "، و" التفسير" ، و" الحجم ، و" الترتيب"، و"مثال " ، و"مدى أو سعة القيمة value space " ، "ونوع البيانات" . مع ملاحظة أن خاصية " الحجم " تحدد عدد القيم التي يسمح للعنصر أن يأخذها ، ودائمًا ما تشغل العناصر غير المتكررة بالنسبة لخاصية "الحجم" بالقيمة " ١ " ، أما بالنسبة للعناصر المتكررة ، فتحدد خاصية الحجم أصغر عدد يطلب من النظام المتوافق أن يدعمه، وتشير خاصية " الترتيب " إلى مدى وجود نظام معين لترتيب العناصر المتكررة ، كالترتيب من العام إلى الخاص مثلاً . أما خاصية " سعة القيمة " فتستخدم لتحديد القيم المسموح باستخدامها ، سواء عن طريق الرجوع إلى قائمة استنادية معتمدة من قبل

(١) Carl Lagoze, "Business Unusual: How 'Event -Awareness' May Breathe Life into the Catalog;' in Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium, Library of Congress, November ١٥-١٧, ٢٠٠٠, available at http://wiley.ed.usu.edu/docs/encyc.pdf. Accessed ١ July ٢٠٠٢.

مبادرة الكيان التعليمي ، أو بالرجوع إلى خطة خارجية ، وفي حالة استخدام لغة مقيدة ينبغي أن تعطى قيمة العنصر بطريقة " المصدر، القيمة Source,value " ، على أن يشير المصدر إلى اللغة المقيدة ، هكذا:

"LOMv١,٠", "Questionnaire"

أما عنصر " نوع البيانات " فيبين طبيعة القيم ، التي قد تكون واحدة مما يلي : "سلسلة لغوية LangString "، أو " التاريخ والوقت DateTime " ، أو " الفترة Duration "، أو " اللغة المقيدة Vocabulary " ، أو " سلسلة حروف CatacterString "، أو "غير معرف Undefined " . مع ملاحظة أن معظم قيم السلاسل تحدد بالقيمة " سلسلة لغوية " ـ وهي عبارة عن شكل من أشكال " نوع البيانات " يتكون من رمز للغة محصور بين علامتي اقتباس ، متبوعًا بفاصلة، ثم تتبعها القيمة التي تشير للسلسلة اللغوية محصورة كذلك بين علامتي اقتباس ، هكذا :

"en", "١٦th century France"

وعادة تعرف عناصر ما وراء بيانات الكيان التعليمي على أنها أجزاء من تجميعات ، أو من تركيبات تبدأ بالفئة ، وقد تتضمن مستويات وسيطة ، ففي فئة " الخصائص العامة " ـ على سبيل المثال - يشتمل معرف التجميعة على العنصرين : "الفهرس Catalog " (وهو اسم أو محدد خطة تعريف أو فهرسة هذا المدخل ") ، وعنصر " المدخل Entry " (وهو " قيمة المعرف داخل خطة التعريف أو خطة الفهرسة التي تحدد أو تعرف هذا الكيان التعليمي"). ويعرض الشكل رقم (١٢/ ٣) صفحة من مسودة مواصفة ما وراء بيانات الكيان التعليمي لجمعية المهندسين الكهربائيين أو مهندسي

الإلكترونيات ، والتي يتم في سياقها تعريف العناصر الأولى من مجموعة العناصر المتعددة التي تندرج تحت فئة " دورة الحياة .

Name	Explanation	Size	Order	Value Space	Datatype	Example
ife Cycle	This category describes the history and current state of this learning object and those entities that have affected this learning object during its evolution.	1	unspecified	-	-	-
ersion	The edition of this learning object.	1	unspecified	-	LangString (smallest permitted maximum: 50 char)	("en","1.2.alpha"), ("nl","voorlopige versie")
atus	The completion status or condition of this learning object.	1	unspecified	draft final revised unavailable Note: When the status is "unavailable" it means that the learning object itself is not available.	Vocabulary (State)	-
ontribute	Those entities (i.e., people, organizations) that have contributed to the state of this learning object during its life cycle (e.g., creation, edits, publication). Note 1: This data element is different from 3.3:Meta-Metadata.Contribute. Note 2: Contributions should be considered in a very broad sense here, as all actions that affect the state of the learning object.	smallest permitted maximum: 30 items	ordered	-		-

ber

بشكل مكثف على ما وراء البيانات التعليمية ، وبأحد عشر عنصرًا من العناصر الثمانية والخمسين المعرفة الواقعة في فئة " الجوانب التعليمية " ، كما أن هناك عنصرين معنيان بدرجة التفاعلية : أولهما عنصر "مدى التفاعلية interactivity " ، وتستخدم معه القيم المحددة التالية : " نشط " ، " وتفسيري " ، و"مختلط " ، وثانيهما عنصر " مستوى التفاعل Interactivity Level " ، وتتراوح القيم في هذا العنصر من مستوى " منخفض جدًا " إلى " مرتفع جدًا " . ويلاحظ كذلك أن عنصر " نوع مصدر التعلم " يأتي مشابهًا لعنصر " نوع مصدر جيم " ، لكنه يستخدم قائمة استناد مختلفة . كذلك يحاول عنصر " الكثافة الدلالية Semantic Density " تقدير درجة

كثافة أو تركيز كيان التعلم ، وذلك من خلال القيم التي يأخذها ، والتي تتراوح من" منخفض جدًا" إلى " مرتفع جدًا . أما عنصر " الدور المقصود للمستفيد النهائي Intended End User Role " فيشير إلى هوية من يعد الكيان التعليمي لخدمتهم ، هل هم المدرسون ، أم المؤلفون ، أم المتعلمون ، أم المديرون . ويبين عنصر " السياق " طبيعة بيئة التعلم المقصودة ، وهل هي مدرسة، أم تعليم عالٍ ، أم بيئة تدريبية ، أو غير ذلك .

وهناك عناصر أخرى في فئة " الجوانب التعليمية " ، إلا أنها واضحة ولا تحتاج لمزيد من التوضيح ، منها : عنصر " المدى العمري المثالي " ، و" درجة الصعوبة" ، و" الوقت المثالي للتعلم" ، و" الوصف" ، و" اللغة" . وتجدر الإشارة في هذا السياق إلى أن ما وراء البيانات التعليمية والتربوية يمكن وضعها ضمن فئة "التصنيف" إذا كان من الممكن قياس القيمة بمقياس مقنن ، ومن أمثلة ذلك "خطط العمر القرائي"، و"خطط مستوى الذكاء" ، و"المهارات التي يطمح أن يكتسبها المستفيد" ، و"المهام التي ينبغي عليه إنجازها".

ومع أن مواصفة ما وراء بيانات الكيان التعليمي ليس لها قواعد محتوى مرتبطة بها ، فإنها تولي استخدام اللغات المقيدة أهمية كبرى ، والدليل على ذلك أن أكثر من نصف العناصر تشير إلى عملية الاستناد ، يضاف إلى ذلك أن المواصفة نفسها عرّفت بسبع عشرة قائمة استناد ، كذلك تسمح ما وراء بيانات الكيان التعليمي باستخدام لغات أخرى غير متعارضة معها ، ولكنها تهدف من وراء ذلك إلى رفع معدلات عملية التشغيل البيني ، وفي حالة استخدام لغة مقيدة غير معتمدة من قبل ما وراء بيانات الكيان التعليمي ، وكانت القيمة المأخوذة من تلك اللغة فيها أيضاً، فينبغي في هذه الحالة أن تحدد القيمة على أنها مأخوذة عن

ما وراء بيانات الكيان التعليمي .

وعلى عكس التأكيد على استخدام اللغات المقيدة بالنسبة لعناصر الموضوع والتصنيف والعناصر التعليمية ، فإن ما وراء بيانات الكيان التعليمي شأنها في ذلك شأن جيم ، ليس بها ما يلزم وجوب الاعتماد على قوائم استنادية محددة بالنسبة للمؤلفين والمشاركين ، ومع ذلك تنص على إدخال قيم المشاركين طبقًا لمواصفة " vCard " التي تصف بنية معلومات الاسم ومعلومات التواصل معه ، دون اختيار أو صياغة قيم هذه العناصر(١)، إلا أن ذلك قد يعكس تصورًا بأن عملية التأليف تعد أقل أهمية بالنسبة لمواد المناهج التعليمية وكيانات التعلم منها بالنسبة للأعمال الأكاديمية والبحثية .

ومع أن ما وراء بيانات الكيان التعليمي نفسها عبارة عن تركيبة مستقلة بذاتها ، إلا أن ائتلاف التعليم الكوني قد نشر تعريفات خطة لغة الترميز الموسعة XML ، ومعرفات نوع الوثيقة DTDs الخاصة بما وراء بيانات مصدر التعلم ، كما نشرت مبادرة التعليم الموزع المتقدم تعريفات خطة لغة الترميز الموسعة XML الخاصة بامتدادات النموذج المرجعي لكيان المحتوى القابل للمشاركة SCROM إلى خطة ائتلاف التعليم الكوني ويعرض الشكل (٤/١٢) بداية نموذج لتسجيلة ما وراء بيانات في صيغة لغة الترميز الموسعة ، مقتبسة من موقع مبادرة ائتلاف التعليم الكوني.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XML Spy v3.5 (http://www.xmlspy.com) by Boyd N Nielsen (NETg) -->
<lom xmlns="http://www.imsglobal.org/xsd/imsmd_v1p2"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.imsglobal.org/xsd/imsmd_v1p2 imsmd_v1p2p2.xsd">
      <general>
            <identifier>x-ims-plirid-v0.30NS.05-107-9989.nloid.en_US 72475</identifier>
            <title>
                  <langstring xml:lang="en-US">Microsoft SQL Server 7.0: Implementing a Database - Part 1</langstring>
            </title>
            <catalogentry>
                  <catalog>http://www.netg.com/catalog1.html</catalog>
                  <entry>
                        <langstring xml:lang="en-US">72475</langstring>
                  </entry>
            </catalogentry>
            <language>en</language>
            <description>
                  <langstring xml:lang="en-US">This is the first course in a five part series that will provide students with the knowledge to implement a database solution with Microsoft SQL Server 7.0.</langstring>
            </description>
            <keyword>
                  <langstring xml:lang="en-US">Windows NT</langstring>
            </keyword>
            <keyword>
                  <langstring xml:lang="en-US">Microsoft</langstring>
            </keyword>
            <keyword>
                  <langstring xml:lang="en-US">Database</langstring>
            </keyword>
            <keyword>
```

(١) T. I
Info
Dav
http

الشكل رقم ٤/١٢
بداية تسجيلة ما وراء بيانات بسيطة .

الشكل رقم (٤/١٢) بداية تسجيلة ما وراء بيانات بسيطة

***المصدر : موقع ائتلاف التعليم الكوني* IMS site**

وهناك عدة مشروعات تقوم الآن بتجربة استخدام ما وراء بيانات الكيان التعليمي أو خطط أخرى ذات صلة بها، غير أن هذه الخطط معقدة إلى حد ما، ولكي تكون التسجيلات ذات معنى يجب أن تأتي مطولة نسبياً ؛ ومن ثم حدد ائتلاف التعلم الكوني مجموعة أساسية مكونة من تسعة عشر عنصرًا مقترحاً (مع أنها ليست عناصر إجبارية). ولذلك فإن توفير الوقت والمستوى المهاري المطلوبين لإنشاء ما وراء بيانات متوافقة تعد من القضايا الحيوية المثارة والجديرة بالمناقشة، ويزيد الأمر حدة ذلك الافتراض الفلسفي والعلمي القائل بأن كيانات التعلم يجب أن تكون في أصغر حجم

ممكن مما يتطلب معه إنشاء عدد كبير نسبياً من تسجيلات ما وراء البيانات التي تصف العديد من الوحدات الصغيرة. ولكن مع تطوير أدوات حاسوبية ودعم مركزي من مشروعات مدعومة أغلبها بمنح مالية من الولاية أو منح فيدرالية ، فقد بات من السهل إنشاء ما وراء بيانات تعليمية ، يضاف إلى ذلك أن الكثير منها أيضاً احتل مكاناً في سياق البيئات الربحية للشركات المسوقة للقطاع التدريبي الأكاديمي والتجاري. ومن المذهل حقاً ما يحظى به تطوير خطط ما وراء البيانات من تنسيق بين المبادرات التعليمية الكبرى مثل ائتلاف التعليم الكوني ، ولجنة معايير تقنيات التعليم التابعة لجمعية المهندسين الكهربائيين ومهندسي الإلكترونيات، وأدريان ، و مبادرة التعليم الموزع المتقدم لتطوير النموذج المرجعي لكيان المحتوى القابل للمشاركة .

ومع ذلك ، فإن وصف ما وراء البيانات القابلة للتشغيل البيني ما هي إلا جانب محدود من القضية الكبرى لإنشاء كيانات تعلم قابلة لإعادة الاستخدام وقابلة للتشغيل البيني على نطاق واسع. فهناك تحديات أخرى باقية ، منها تناغم كل من المصطلحات ومبادئ التصميم في رحلة تطوير كيانات التعلم نفسها، ناهيك عن المعماريات الفنية الطيعة overarching architectures المناسبة، ونظريات التصميم التعليمي. ومن المتوقع أن نرى استثمارًا قويًا قد يستمر في جميع هذه المجالات على مدار هذا العقد وخاصة في ظل ما نشهده من قوة احتمالات تغيير تقنية كيانات التعلم لمجريات خطى التعليم.

READINGS :

- Specifications discussed in this chapter are cited here, followed by the parent website. The websites should be consulted for current

versions of the specifications, some of which are updated fairly frequently, and for related specifications, tools, and documentation.

- ARIADNE Educational Metadata Recommendation, version ٣,٠, available at http://ariadne.unil.ch/Metadata/. The ARIADNE home page is available at http://www.ariadne-eu.org/.

  The ARIADNE project was discontinued in ٢٠٠٠, but the work is being carried on by the ARIADNE Foundation.

- GEM ٢,٠, available at http://www.geminfo.org/Workbench/gem٢.html. The home page of the GEM Project Site for project participants is available at http://www.geminfo.org/. The Gateway to Educational Materials, the GEM search interface, is available at http://www.thegateway.org/welcome.html.

- Greenberg, Jane, ed. Metadata and Organizing Educational Resources on the Internet. New York: Haworth Press, ٢٠٠٠. Simultaneously published as Journal of Internet Cataloging ٣, nos. ١ and ٢/٣ (٢٠٠٠).

  A collection of essays covering several different metadata schemes, projects, and applications.

- IEEE Draft Standard for Learning Object Metadata (IEEE PI٤٨٤,١٢,١/D٦,٤), available at http://ltsc.ieee.org/doclwgI٢/LOM٣_٠٠.pdf. The home page of the IEEE Learning Object Standards Committee, IEEE P١٤٨٤,١٢ Learning Object Metadata Working Group is avail- able at http://ltsc.ieee.org/wgI٢/.

- IMS Learning Resource Meta Data Information Model, version ١,٢ Final Specification, available at http://www.imsproject.org/metadata/imsmdvip٢p l/imsmd_infovl p٢pl.html. The home page of the IMS Global Learning Consortium is available at http://www.imsproject.org/.
- Sharable Content Object Reference Model (SCORM) ١,٢ Content Aggregation Model, available at http://www.adlnet.org/ADLDOCS/Document/SCORMl.٢_CAM.doc. The home page of the Advanced Distributed Learning Network is available at http://www.adlnet.org/.
- Wiley, David A., ed. The Instructional Use of Learning Objects. Association for Instructional Technology, ٢٠٠١. Online version available at http://www.reusability.org/read/.

  A collection of essays that together constitute a primer on the theory of learning objects and the practice of implementing them. Although this is not the focus, some of the essays touch on metadata standards and issues. The entire book is available online, with the ability to comment on and submit corrections to the text.

الفصل الرابع عشر

ما وراء البيانات للمصادر الجغرافية المكانية والبيئية

ما وراء البيانات للمصادر الجغرافية المكانية(١) والبيئية :

من اللافت للانتباه أن كثيراً من مصادر المعلومات تشتمل على بعض المكونات الجغرافية ، فهنالك من الكتب والمقالات التي كتبت عن موضوعات تعود جذورها إلى مكان معين، مثال ذلك أحد الكتب التي تتناول تاريخ دولة البيلاروس Belarus ، أو أحد الكتب الأخرى التي تتناول حانات البيانو في نيويورك ،كما أن الصحف اليومية عادة ما تميل إلى معالجة قضايا ذات أبعاد إقليمية ، بل إن الصور الفوتوغرافية عادة ما تمثل أماكن معروفة ، كما لا يخفى أن يكون لمجموعات البيانات في مجال العلوم الاجتماعية والبحوث العلمية أبعادها الجغرافية . من هذا المنطلق حرصت معظم خطط ما وراء البيانات في وصفها لمصادر المعلومات على تخصيص بعض العناصر لتسجيل المعلومات الجغرافية .

ونستهل هذا الفصل بالتعريف بأحد المعايير التي صممت لوصف المصادر الجغرافية المكانية geospatial الرقمية ، ثم نعرج بعد ذلك للحديث عن بعض خطط ما وراء البيانات المخصصة لوصف مصادر المعلومات البيولوجية والنوعية التي غالباً ما يكون لها مكوناتها ذات الأبعاد المكانية .

معيار محتوى ما وراء البيانات المكانية **FGDC**:

معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية The Content

(١) يقصد بالمعلومات الجغرافية المكانية Geospatial المعلومات المرجعية وغير المرجعية للأماكن الجغرافية ، وهي بذلك أشمل من المعلومات ذات الطابع الجغرافي والتي تهتم بالجوانب الطبيعية والموقع والسكان وغيرها [المترجم].

Standard for Digital Geospatial Metadata CSDGM أحد أقدم المواصفات التي وسمت نفسها بأنها معيار لما وراء بيانات ، وقد طورت هذا المعيار اللجنة الفيدرالية للبيانات الجغرافية(١)، في بدايات عام ١٩٩٢م، لكن أولى إصدارات هذا المعيار صدرت عام ١٩٩٤م، ولقد اكتسب هذا المعيار صفته الرسمية عندما قضى أحد القرارات الإدارية - رقم ١٢٩٠٦ - والموقع من قبل الرئيس الأمريكي [السابق] كلينتون ، وعنوانه " تنسيق اقتناء البيانات الجغرافية وإتاحة الوصول إليها : البيئة التحتية للبيانات المكانية الوطنية Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure " ، بأن تستخدم الهيئات الفيدرالية التي تقتني بيانات ذات طبيعة جغرافية مكانية أو تتيح الجديد منها معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية في وصف تلك البيانات ، كما يجب عليها أن تتيح ما يتوافر لديها من بيانات لأحد المراكز الوطنية لتبادل البيانات المكانية National Geospatial Data Clearinghouse ، على أن يقوم هذا المركز والذي ترعى شؤونه اللجنة الفيدرالية للبيانات الجغرافية ، بتقديم خدمات البحث الفيدرالي لما وراء البيانات المكانية من خلال مستودعات لا مركزية لما وراء البيانات اعتماداً على المعيار Z٣٩,٥٠ ، وبأسلوب أشبه ما يكون بالأسلوب الذي يتبعه مكتب الولايات المتحدة للمطبوعات الحكومية GPO في تقديمه لخدمات بحث ما وراء بيانات جيلس GILS.

وتجدر الإشارة إلى أن البيانات الجغرافية المكانية الرقمية يتسع مجالها

(١) لجنة مكونة من ممثلين من سبع عشرة هيئة اتحادية (فيدرالية) لتشجيع الاستخدام الوطني للبيانات الجغرافية [المترجم].

لتشمل الكثير من البيانات بدءاً من الصور التي يتم التقاطها عن طريق الأقمار الصناعية حتى مجموعات بيانات نظم المعلومات الجغرافية، وبناءً على معيار محتوى ما وراء البيانات المكانية الرقمية، تتمثل الوظائف التي تقوم بها ما وراء البيانات المكانية الرقمية في الوظائف الأربعة التالية :

- تحديد البيانات التي تتوافر عن أحد المواقع الجغرافية .
- التثبت من مدى تلبية البيانات لحاجات معينة (مدى الاستفادة منها).
- إتاحة الفرصة للمستفيد للحصول على مجموعة بيانات محددة .
- تيسير معالجة مجموعة البيانات والإفادة منها .

وفي يونيو من عام ١٩٩٨م نشرت الإصدارة الحالية من معيار محتوى ما وراء البيانات المكانية الرقمية ، وتتاح عبر موقع اللجنة الفيدرالية للبيانات الجغرافية على الإنترنت (http://www.fgdc.gov/metadat/constan.html). وتعرف هذه الإصدارة بكل من عناصر البيانات ، والعناصر المركبة compound elements - أي العناصر التي تتألف من مجموعات من عناصر البيانات وبعض العناصر المركبة الأخرى ، أو تعرف بكليهما ، ويتم وصف عناصر البيانات من خلال ست خصائص ، هي : الاسم ، والتعريف ، والنوع ، والنطاق أو المجال Domain ، والاسم المختصر Short name ؛ حيث تشير خاصية " النوع إلى نوعية البيانات التي تعطي للقيمة ، مثال ذلك : عدد صحيح ، ورقم حقيقي ، نص ، تاريخ، وقت أو زمن . أما خاصية " المجال " فيتم من خلالها تحديد القيم التي تعطى لأحد العناصر ، إما في شكل قائمة بالقيم ، أو عن طريق الإشارة إلى إحدى القوائم الاستنادية الخارجية ، أو كمجموعة من

القيود التي تحدد القيم التي يمكن أن تخصص للعنصر . أما خاصية " الاسم المختصر " فهي عبارة عن اسم مكون من ثماني تمثيلات أو أقل من ذلك ويستخدم لتيسير تطبيق المعيار ؛ فالاسم المختصر مثلا لعنصر "مرات الصيانة والتحديث " هو" update " ، وعلى الطرف الآخر تعرف العناصر المركبة من خلال ثلاث خصائص هي: الاسم ، والنوع (وهو عادة ما يكون مصطلحا مركباً compound ") ، والاسم المختصر ، كما يمكن أن توصف هذه العناصر المركبة من خلال مجموعة من " قواعد الإنتاج production rules "، التي تبين توليفة العناصر والمركبات التي تتألف منها . ويعرض الشكل رقم (١٤/ ١) إحدى صفحات معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية التي تصف مركب " المجال أو النطاق المكاني".

تنظيم عناصر بيانات معيار محتوى ما وراء البيانات المكانية الرقمية

جاء هذا المعيار في عشرة أقسام رتبت على النحو التالي :

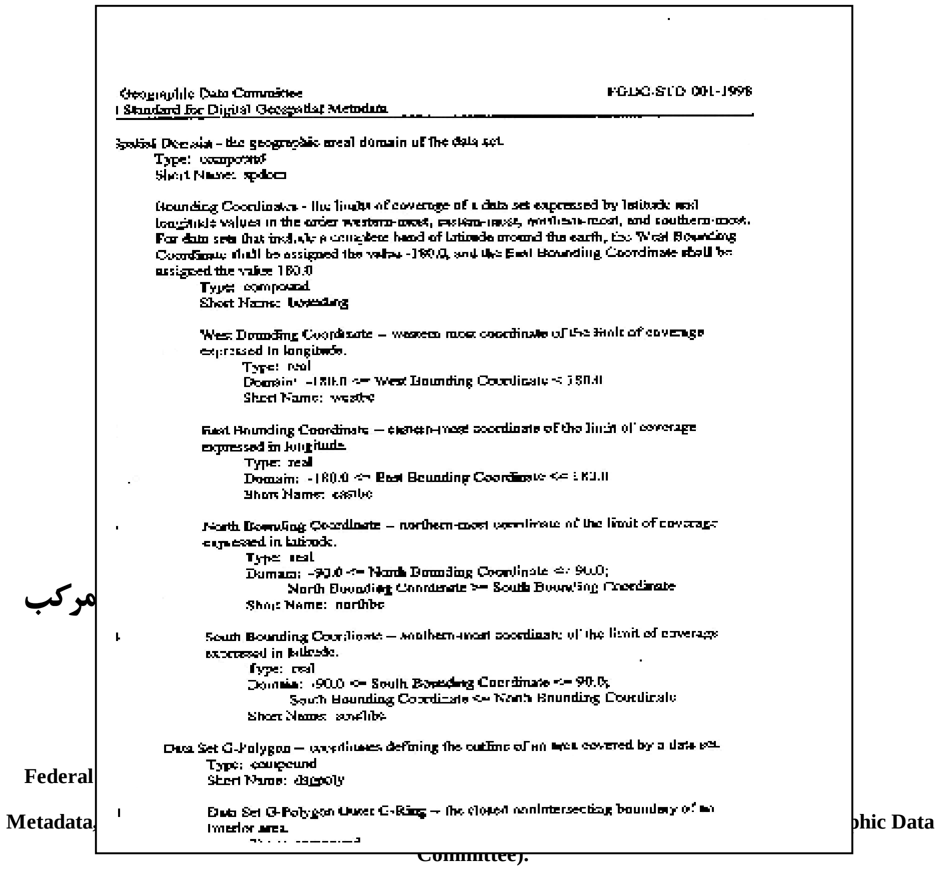
- معلومات التعريف .
- معلومات دقة البيانات .
- معلومات تنظيم البيانات المكانية .
- المعلومات المرجعية المكانية .
- معلومات الكيان والخاصية .
- معلومات التوزيع .
- المعلومات المرجعية لما وراء البيانات .

- معلومات الإرجاعة الببليوجرافية
- معلومات الفترة الزمنية .
- معلومات الاتصال .

ومن الملاحظ أن الأقسام السبعة الأولى تشكل تسجيلة ما وراء البيانات ، بينما تعرف الأقسام الثلاثة الأخيرة ببيانات المعلومات العامة التي تستخدم في قسمين أو أكثر من أقسام ما وراء البيانات السابقة .

وتجدر الإشارة في هذا الصدد إلى أن هناك قسمين من المعلومات يجب تضمينهما إجبارياً في تسجيلة ما وراء البيانات ، هما : معلومات التعريف ، والمعلومات المرجعية لما وراء البيانات ، أما بقية الأقسام الأخرى فتسجل بشكل إجباري فقط عندما يتطلب الأمر ذلك . ومن الملاحظ كذلك أن قسم معلومات التعريف يتضمن عنصراً مركباً إجبارياً وهو العنصر الخاص بمعلومات الإرجاعة الببليوجرافية ، التي تُعّرف في قسم مستقل هو قسم معلومات الإرجاعة الببليوجرافية ، ومعلومات الإرجاعة هذه في المعيار يمكن أن تشتمل على معلومات ببليوجرافية، مثل : العنوان، واسم المنشئ creator (الذي يسمى في سياق هذا المعيار originator)، والطبعة، ومكان النشر ، واسم الناشر ، وتاريخ النشر، وبيانات السلسلة، كما أن هذه الفئة من المعلومات يمكن أن تشتمل على بيان خاص بوقت النشر إذا كان معروفاً ، وعلى عنصر آخر يسمى " نموذج عرض البيانات الجغرافية المكانية Geospatial Data Presentation Form " وهو من العناصر التي يجب أن يعتمد في اختيارها على إحدى القوائم الاستنادية للمصطلحات المناسبة مثل :

خريطة، وكرة أرضية، ونموذج، وبيانات رقمية خطية raster dd، وبيانات رقمية شبكية vector dd، وصورة مأخوذة عن طريق الاستشعار عن بعد. وبالإضافة إلى معلومات الإرجاعة البيليوجرافية، يشتمل قسم معلومات التعريف على مركبات إجبارية خاصة لكل من : الوصف، والفترة الزمنية للمحتوى، والحالة، والنطاق المكاني، والكلمات المفتاحية، كما يتضمن



Geographic Data Committee — FGDC-STD-001-1998
Standard for Digital Geospatial Metadata

Spatial Domain – the geographic areal domain of the data set.
Type: compound
Short Name: spdom

Bounding Coordinates - the limits of coverage of a data set expressed by latitude and longitude values in the order western-most, eastern-most, northern-most, and southern-most. For data sets that include a complete band of latitude around the earth, the West Bounding Coordinate shall be assigned the value -180.0, and the East Bounding Coordinate shall be assigned the value 180.0
Type: compound
Short Name: bounding

West Bounding Coordinate – western-most coordinate of the limit of coverage expressed in longitude.
Type: real
Domain: -180.0 <= West Bounding Coordinate < 180.0
Short Name: westbc

East Bounding Coordinate – eastern-most coordinate of the limit of coverage expressed in longitude.
Type: real
Domain: -180.0 <= East Bounding Coordinate <= 180.0
Short Name: eastbc

North Bounding Coordinate – northern-most coordinate of the limit of coverage expressed in latitude.
Type: real
Domain: -90.0 <= North Bounding Coordinate <= 90.0;
North Bounding Coordinate >= South Bounding Coordinate
Short Name: northbc

South Bounding Coordinate – southern-most coordinate of the limit of coverage expressed in latitude.
Type: real
Domain: -90.0 <= South Bounding Coordinate <= 90.0;
South Bounding Coordinate <= North Bounding Coordinate
Short Name: southbc

Data Set G-Polygon – coordinates defining the outline of an area covered by a data set.
Type: compound
Short Name: dsgpoly

Data Set G-Polygon Outer G-Ring – the closed nonintersecting boundary of an interior area.

مركب

Federal

Metadata, ... hic Data Committee).

كذلك عناصر إجبارية خاصة بضوابط الإتاحة أو الوصول، أو ضوابط الاستخدام . مع ملاحظة أن المركب المسمى " الوصف " ينبغي أن يشتمل

علـى مستخلص وجملة نصية تعبر عن الهدف يصفان الدافع من وراء إنشاء مجمـوعة البـيانات. أمـا مـركب " الحالـة " فيـشتمل علــى عنـصر مـرات التحديث، والحالـة التـي يشار من خلالها إلى حالة البيانات ، بمعنى هل هي مكتملة أم فـي طور الإعداد، أم لا تزال فـي مرحلة التخطيط . أمـا مركب "الكلمـات المفتاحية أو الأساسية " فيقسم إلى أربعة عناصر هي : الموضوع الرئيـسي (تحـدد موضـوعياً)، والمكـان (الموقـع الجغرافـي)، والطـبقة Stratum (الموقـع الرأسـي)، والـبعد الزمنــي Temporal ، مـع ملاحظــة أن المصطلحات التـي تـرد فـي سياق هذه العناصر قد يعتمد في صياغتها على لغات مقيدة، وقد تكون مصطلحات غير مقيدة، على أن يتضمن كل مركب مـن المـركبات المسماة " كلمـة مفتاحـية " عنـصراً يـشير إلـى المكنـز الـذي أخذت منه الكلمة المفتاحية وعنصراً آخر للكلمة المفتاحية نفسها.

أمـا عنـصر "الفتـرة الزمنية للمحتوى" فيتم في سياقه تمثيل الفترة تمثيلاً يعبـر عنه تاريخًـا أو زمنًا واحدًا ، أو عن عدة تـواريخ أو أزمنة ، أو مدى تاريخي أو زمني محدد. أما مركب "النطاق المكاني" فيشتمل على إحداثيات إجبارية للحدود الشمالية والجنوبية والشرقية والغربية bounding coordinates يـتم التعبيـر عـنها بقيم خطـوط الطـول والعـرض ، كمـا يمكـن أن يمـثل هـذا المركب _ أي النطاق المكاني _ اختيارياً في شكل مضلع جي G-Polygon ، الـذي يتم من خلاله وصف المنطقة عن طريق مجموعة من إحداثيات تسمى حلقات جي G-rings.

ويشتمل قسم "معلومات دقة البيانات" على معلومات أقرب ما تكون إلى تقيـيم مـدى صـحة قـيم الخـصائص فـي مجمـوعة البـيانات، أي تقيـيم لمـدى

التوافق المنطقي للمعلومات التي تتضمنها مجموعة البيانات، بما في ذلك أي اختبارات أجريت للبيانات، ثم يتضمن كذلك معلومات عن مدى اكتمال البيانات، والمعلومات التي تؤصل مصدر البيانات والأطراف المسؤولة عنه lineage information.

أما "معلومات تنظيم البيانات الجغرافية المكانية" فتصف نموذج البيانات المكانية المستخدم في ترميز هذه البيانات، ويتبع هذا القسم قسمان فرعيان لكل من "المرجع المكاني المباشر direct spatial reference " و"المرجع المكاني غير المباشر indirect spatial reference " ؛ حيث يشتمل "المرجع غير المباشر" على بيانات بأسماء الخصائص الجغرافية ، وبيانات تقرر الخطط وغيرها من الوسائل الأخرى التي يتم من خلالها الإحالة إلى المواقع، مثال ذلك أسماء الشوارع . أما "المرجع المكاني المباشر"، فينبغي أن تكون إما في شكل نقطة، أو خطاً vector (خطوط موجهة)، أو شبكة raster (خطوط شبكية grid)، مع ملاحظة أن المعلومات في شكل النقطة والخط تعتمد في التعبير عنها على "مفاهيم البيانات المكانية Spatial Data Concepts " التي يتضمنها " معيار نقل البيانات المكانية لوزارة التجارة the Department of Commerce Spatial Data Transfer Standard "، أو يعبر عنها بمصطلحات مأخوذة من " صيغة منتجات الخطط الموجهة لوزارة الدفاع the Department of Defense Vector Product Format ".

كذلك يتكفل قسم "المعلومات المرجعية المكانية" بالإشارة إلى الإحداثيات الأفقية والرأسية (أي الارتفاع والعمق) في مجموعة البيانات؛ فإذا كانت خطوط الطول والعرض - على سبيل المثال - مستخدمة

كإحداثيات أفقية ، فيجب في هذه الحالة تحديد كل من فارق خطوط الطول والعرض (أي الحد الأدنى الفارق بين قيمتين متجاورتين)، ونوع الوحدات المستخدمة للقيم (مثال ذلك " درجات عشرية decimal degrees") .

أما قسم "معلومات الكيان والخاصية" فيحدد طبيعة المعلومات الجغرافية المتضمنة [في المصدر]، مثال ذلك الطرق، أو الخصائص الجغرافية، أو المرتفعات، كما يحدد كذلك كيفية تمثيل هذه البيانات . وقد يتضمن هذا القسم وصفاً تفصيلياً لكل خاصية، أو مجرد عرض وصفي مختصر مع روابط إلى وصف أكثر اكتمالاً خارج عن حدود تسجيلة ما وراء البيانات . ومن الملاحظ أن قيمة هذه الإمكانية الأخيرة تكون أكثر إفادة في الحالات التي يتوافر فيها لمجموعة البيانات توثيق خارجي جيد، ولا يستلزم الأمر إعادة تكراره في تسجيلة ما وراء البيانات .

أما "معلومات التوزيع " فتشير إلى المكان الذي يتم من خلاله الحصول على البيانات ، وإلى نوعية الصيغ المتوافرة . أما قسم "المعلومات المرجعية لما وراء البيانات metadata reference information " فهو بمثابة بيانات وصفية meta-metadata لتسجيلة ما وراء البيانات نفسها ، حيث يتضمن معلومات عن التاريخ والوقت اللذين أنشئت فيهما التسجيلة، وتاريخ التحديث الأخير لها ، والوقت الذي ينبغي أن تراجع فيه، ورقم إصدارة معيار محتوى ما وراء البيانات المكانية الرقمية المستخدمة، والشخص المسؤول الذي يمكن الاتصال به لمزيد من المعلومات.

قواعد المحتوى والتركيبة :

إذا كان معيار محتوى ما وراء البيانات المكانية الرقمية قد حدد صيغاً لعناصر التاريخ، والوقت، وقيم الطول والعرض ، فمن الملاحظ أنه لم يشتمل على أية قواعد محددة لصياغة المحتوى ، إذ تأخذ أسماء المنشئين وهيئات الاتصال قيماً نصية حرة غير استنادية أو دون الالتزام بصيغ مفضلة للأسماء ، ولم تأت إشارة المعيار إلى الاعتماد على القوائم الاستنادية المتداولة في عالم المكتبات إلا في بضعة مواضع ، منها أن القائمة الاستنادية لنموذج عرض البيانات الجغرافية المكانية تعتمد في جزء كبير منها على " الدليل الإرشادي لفهرسة المواد الخرائطية في ضوء القواعد الأنجلو أمريكية للفهرسة" الموسوم "Cartographic Materials: A Manual of Interpretation for AACR٢" (Chicago: American Library Association, ١٩٨٢). كما أن تعريف " مركب الكلمة المفتاحية للموضوع " يشير إلى الاعتماد على قائمة المكانز التي تتضمنها قائمة رموز مارك للروابط والمصادر الموسومة "MARC Code List for Relators, Sources, and Description Conventions" (Washigton,D.C.: Library of Congress, ١٩٨٨). أما بالنسبة " لمكنز الكلمة المفتاحية للمكان " فقد أدرج نظام معلومات الأسماء الجغرافية كقيمة استنادية ، في إشارة إلى أن هذا المكنز هو المفضل في هذا السياق ، أو على الأقل بما يوحي بأنه هو المتوقع استخدامه على نطاق واسع.

ويلاحظ في هذا الصدد أن معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية لا يحتاج إلى أي تركيبة نقل transport syntax محددة ، إلا أن اللجنة الفيدرالية للبيانات الجغرافية تتيح عبر موقعها على الإنترنت

معرفاً معتمداً من قبلها من معرفات نوع الوثيقة في صيغة لغة الترميز الموسعة XML DTD ، بالإضافة إلى بعض نماذج عرض الصفحات التي تستخدم في عرض البيانات اعتماداً على هذا المعرف .

الامتدادات والبروفايلات :

يلحق بالإصدارة المنقحة الصادرة عام ١٩٩٨م من معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية ملحقان ، يختص أولهما بقواعد التوسعة أو الامتدادات ، ويختص الآخر بقواعد إنشاء البروفايلات. وقد عُرّفت هذه القواعد تعريفاً دقيقاً حتى تأتي متوافقة مع بنية المعيار وتقنياته . مع ملاحظة أن الامتدادات Extensions ينبغي أن توثق بشكل رسمي ويشار إليها في مركب "امتدادات ما وراء البيانات Metadata Extensions " الذي يرد ضمن قسم "المعلومات المرجعية لما وراء البيانات " . كما أن تعريف العناصر والمركبات الموسعة ينبغي أن تشتمل على الخصائص المستخدمة في تعريفات هذا المعيار مع أربع خصائص إضافية هي: خاصية " المصدر " (أي اسم الكيان الذي أنشأ العنصر)، وخاصية " السبب الجوهري rational " (أي السبب الاختياري لإنشاء العنصر)،وخاصية "الأصل Parent " (أي العنصر الموجود في المعيار والذي يظهر تحته العنصر الجديد) ، وخاصية "الفرع Child " (أي العناصر التي يمكن أن تظهر تحت العنصر الجديد) .

أما البروفايلات فهي عبارة عن تهيئة للمعيار لاستخدامات مجتمع معين، إذ تخصص البروفايلات المجموعة الفرعية لعناصر المعيار التي سيستخدمها ذلك المجتمع، كما قد تشتمل على امتدادات معرفة بشكل رسمي

. ومن الملاحظ أن البروفايل قد لا يمكنه أن يغير تعريف أو استخدام أي من العناصر الموجودة، ولكن يجب كذلك أن يشتمل على جميع العناصر الإجبارية، والعناصر الإجبارية عندما يتطلب الأمر ذلك، إلا أنه يمكنه مع ذلك تحويل صفة بعض العناصر الاختيارية إلى عناصر إجبارية ، وتقييد قيم المجالات الخاصة ببعض العناصر . ومن الملاحظ أن هذه البروفايلات يمكن أن تستخدم بشكل رسمي إذا ما أقرت من قبل اللجنة الفيدرالية للبيانات الجغرافية ، أو أن تستخدم بشكل غير رسمي من قبل أحد مجتمعات المستفيدين، ولذلك فإن اللجنة تقوم بإعداد قوائم بالبروفايلات التي أقرتها ، أو التي لم تقرها بعد ولا تزال في طور الإعداد عبر موقعها على الإنترنت (http://www.fgdc.gov/metadata/constant.html).

ويمكن القول إن معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية من المعايير التي يكتنف استخدامها قدر كبير من الصعوبة ؛ فالمعيار نفسه كبير جداً، كما يعرف بأكثر من ثلاثمئة عنصر ومركب، ناهيك عن أن تعقد القواعد التي يتم بناء عليها تحديد العناصر الإجبارية، والعناصر الإجبارية عندما يتطلب الأمر ذلك، والعناصر الاختيارية، يضاف إلى ذلك أنه وبحكم تعدد العناصر الإجبارية التي يضمها هذا المعيار، فإن وصف المصدر التي يعتمد فيه على هذا المعيار قد يأتي مطولاً بشكل كبير ؛ إذ يبلغ طول كل تسجيلة من التسجيلتين اللتين جاءتا كنموذجين في الدليل الإرشادي الخاص بهذا المعيار والموسوم The Content Standard for Digital Geospatial Metadata Workbook " نحو سبع صفحات ؛ ولهذا السبب رأت بعض الولايات أن تعد خططًا مبسطة metadata-lite لما وراء البيانات

الجغرافية المكانية خاصة بها. كذلك تبنى أحد اجتماعات القمة الخاص بما وراء البيانات - والذي عقد في دنفر في فبراير من عام ١٩٩٦م، تعريفًا لمجموعة فرعية من عناصر معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية ، يشار إليها في بعض الأحيان بمجموعة دنفر الأساسية أو دنفر كور Denver Core .

معيار ما وراء البيانات البيولوجية :

يعد بروفايل البيانات البيولوجية المعروف بمعيار ما وراء البيانات البيولوجية للبنية التحتية الوطنية للمعلومات البيولوجية NBII Biological Metadata Standard، أحد أشهر البروفايلات المعروفة لمعيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية، وقد طور هذا البروفايل ضمن جهود مبادرة البنية التحتية الوطنية للمعلومات البيولوجية التي تبنتها مجموعة عمل البيانات البيولوجية الممثلة لكل من اللجنة الفيدرالية للبيانات الجغرافية FGDC، وقسم المصادر البيولوجية للمسح الجيولوجي للولايات المتحدة Biological Resources Division of the U.S Geological Survey (USGS) . وقد أقرت تلك اللجنة الفيدرالية هذا المعيار رسمياً في عام ١٩٩٩م .

والبنية التحتية الوطنية للمعلومات البيولوجية هي برنامج تعاوني تتضافر فيه جهود أطراف يمثلون جهات فيدرالية وحكومية ودولية وغير حكومية وأكاديمية ومن القطاع الخاص، بهدف الارتقاء بسبل إتاحة الوصول إلى البيانات والمعلومات عن المصادر البيولوجية (http://www.nbii.gov) .

وكان جل اهتمام هذا البرنامج منذ بدايته منصرفاً نحو تطوير معيار لما وراء البيانات البيولوجية، ولكن بحكم ما تنطوي عليه النسبة الكبرى من بيانات العلوم البيولوجية والبيئية biological and ecological من مكونات جغرافية مكانية ، فقد رأى القائمون على إعداده أن يبني على معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية ، مع توسعة مجال تطبيقه ليشمل البيانات البيولوجية والبيئية .

وترتكز توسعات بروفايل البيانات البيولوجية على ثلاثة مجالات رئيسية من المجالات التي يغطيها معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية، يتمثل أولها في توسعة قسم "معلومات التعريف" ليشمل معلومات تصنيفية مكثفة taxonomy ، وذلك في المركب المسمى " تصنيف Taxonomy "، ويتمثل ثانيها في إضافة مركب يسمى " أداة تحليلية Analytical Tool " لقسم "معلومات التعريف" كذلك ليتسع مجاله للمعلومات عن جميع الأدوات التحليلية المستخدمة في تحليل البيانات أو يتطلب تفسير البيانات الاعتماد عليها، أما ثالثها فتتمثل في إضافة معلومات خاصة بتوثيق المنهج المستخدم في جمع البيانات ، وذلك في قسم " معلومات دقة البيانات Data Quality Information " . وفضلاً عن ذلك يقوم البروفايل كذلك بتوسعة معلومات المجال الخاصة بالعديد من العناصر الأخرى لتشمل المصطلحات والمصادر البيولوجية المتداولة .

لقد أثار تطوير معيار ما وراء البيانات البيولوجية قضية كبرى تتعلق ببروفايلات معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية ، ذلك

أن بعضاً من مصادر المعلومات البيولوجية كبيانات التجارب المختبرية - على سبيل المثال - ليس لها مكوناتها الجغرافية ، وقد كان لدى المجتمع البيولوجي رغبة في أن يكون مركب " المجال المكاني " في معيار محتوى ما وراء البيانات المكانية الرقمية من المركبات " الإجبارية عندما يكون ذلك ضرورياً " بحيث يمكن استخدامه حسب الحاجة ، وبدون أن تكون إجبارية للمعلومات غير الجغرافية ، ومع ذلك ، وبينما كان يجري تطوير المعيار البيولوجي ، كانت إصدارة عام ١٩٩٨م المنقحة من معيار محتوى ما وراء البيانات المكانية الرقمية قد نشرت وأشير في سياقها إلى وجوب أن تتضمن البروفايلات جميع العناصر الإجبارية الموجودة في ذلك المعيار، وقد وافقت جماعة المعايير التابعة للجنة الفيدرالية للبيانات الجغرافية على إجراء تغيير في قواعد البروفايل ، وبما يسمح للمتطلبات التي يفرضها موقف معين أن تتسع لمجموعات البيانات ذات الطابع غير الجغرافي non geographically-based information ، ومن اللافت للنظر ، أن هذه التجربة أظهرت شكوى عامة من معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية مفادها أن هذا المعيار صمم بشكل متماسك بعضه مع بعض monolithic، ولم يصمم ليستخدم كمكون قابل للتعديل(١) modular ليتوافق مع خطط ما وراء بيانات أخرى .

لغة ما وراء البيانات الإيكولوجية و البيئية :

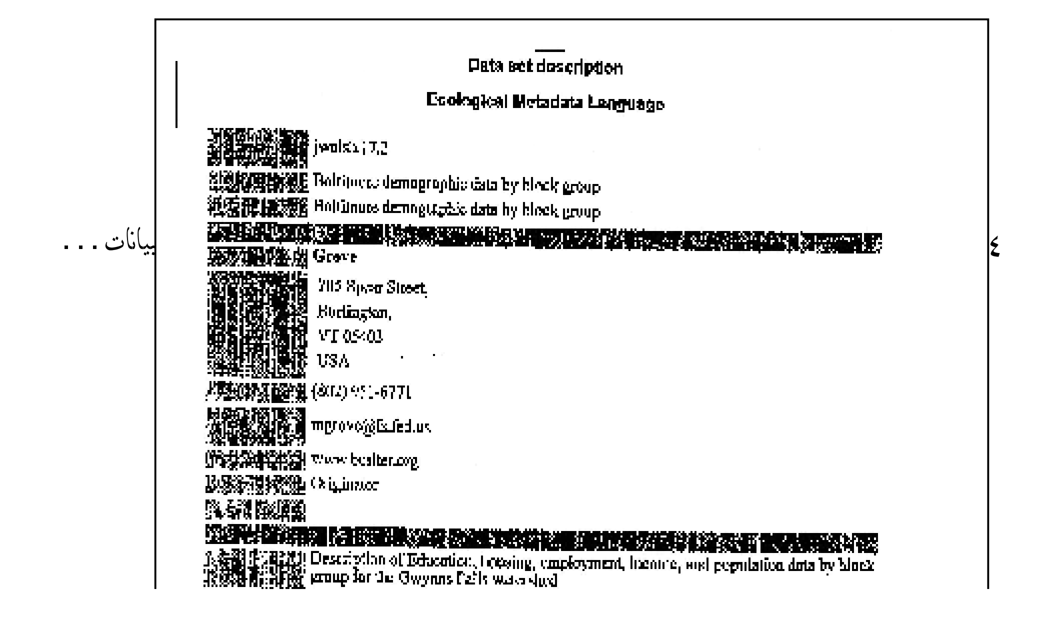
---

(١) يشير مصطلح modular إلى طريقة لتشكيل نظام آلي أو برمجي باستعمال مجموعة من الوحدات المتلائمة ، كما يشير مصطلح modularity إلى قابلية التركيب والتعديل، أي قابلية مكونات نظام آلي أو برمجي للتوسع بإضافة لعض الوحدات الأساسية [المترجم].

في بيئة ما وراء البيانات الحالية ، يمكن أن تجني ثمرات كثيرة من وراء عملية جمع عناصر ما وراء البيانات من عدة خطط مختلفة ، كما هو الحال عند استقاء ما وراء البيانات الوصفية من خطة دبلن كور ، بدلاً من تضمين عناصر وصفية في إحدى خطط المعلومات الجغرافية أو البيئية . وهذا هو التوجه الذي يتبناه مشروع لغة ما وراء البيانات الإيكولوجية the ecological Metadata Language لشبكة المعرفة للمركبات الحيوية Knowledge Network for Biocomplexity (KNB) (http://knb.ecoinformatics.org/) . وقد طُورت هذه الشبكة - التي تتخذ من جامعة كاليفورنيا في سانتا باربرا مقراً لها - بهدف تيسير إعداد البحوث العلمية الإيكولوجية والبيئية عن المركبات الحيوية من خلال توفير سبل الوصول إلى البيانات الموزعة ، وتوفير معلومات استخدامها أو الإفادة منها بفاعلية . ولعل مما يلاحظ في هذا الصدد أن مجموعات البيانات الناتجة عن أسلوب الملاحظة Observation data تتقاسمها بعض السمات، فهي من ناحية أولى موزعة على نطاق واسع بين المؤسسات البحثية والباحثين، وهي من ناحية أخرى غير متجانسة من حيث طبيعتها ومن ثم تتطلب الإفادة منها استخدام أدوات تحليلية تركيبية ؛ من هنا فإن لغة ما وراء البيانات الإيكولوجية هي خطة ما وراء بيانات وصفية قابلة للتركيب modular metadata description scheme في صيغة لغة الترميز الموسعة ، يتم بموجبها تعريف أوصاف السمات العامة لبيانات الملاحظة ، بناء على ذلك يمكن أن يشتمل الوصف المعتمد على هذه اللغة على أية نظم فرعية قابلة للتطبيق

علـى مصدر معلومات بعينه ، كما أنها يمكن أن تشتمل على أجزاء وصفية مأخـوذة مـن خطـط مـا وراء بيانات أخـرى طالمـا أنهـا ممثلـة فـي صيغة لغة الترميز الموسعة .

وتشتمل لغـة مـا وراء البيانات الإيكولوجية نفسها على نظم فرعية ، أو مجموعات عناصـر ما وراء بيانات تختص بوصف كثير من السمات ، مثل : ضوابط الوصول إلى المعلومات ، وقيود استخدامها ، والمعلومات الخاصة بالتغطية الزمنية والمكانية والتصنيفية ، والمعلومات الشخصية ، ومعلومات الإرجاعـة الببليوجـرافية . وانطلاقـاً مـن الافتـراض القائـل بـأن عـددًا مـن مجموعات البيانات الإيكولوجية والبيئة يتم تمثيلها في قواعد بيانات محلية كمجموعات مـن الجداول المترابطة أو العلاقية، سنجد أن هناك عدداً مـن النظم الفرعية لهذه اللغة يعنى بوصف أحد جوانب قاعدة البيانات ، من ذلك إعطـاء نظـرة عامـة علـى المجمـوعات، ووصـف جمـيع الخـصائص (المتغيـرات) التـي يتـضمنها أحـد الجـداول ، وعقـبات الـتكامل البنـيوي structural integrity بين الأعمدة فـي الجداول المختلفة ، كذلك تركز الخطة علـى المعلـومات الخاصـة بالجـوانب المنهجـية ، كمـا هـو الحـال بالنـسبة للعمليات التـي تجرى علـى إحدى مجموعات البيانات للتثبت من مدى دقة بياناتها أو لرفع درجة هذه الدقة .



الشكل رقم (٢/١٤) تسجيلة وصف أحد المصادر بلغة ما وراء البيانات الإيكولوجية EML
مسترجعة من :

نظام بحث شبكة المعرفة للمركبات الحيوية KNB (http://knb.ecoinformatics.org/).

ويعرض الشكل رقم (١٤ / ٢) تسجيلة وصف أحد المصادر بلغة ما وراء البيانات الإيكولوجية مقتبسة من أحد العروض المتاحة مباشرة للمعلومات في نظام البحث الخاص بشبكة المعرفة للمركبات الحيوية .

فئات داروين الأساسية أو داروين كور :

ثمة خطة أخرى معنية بالبيانات الناتجة عن الملاحظة تعرف بداروين كور أو فئات داروين الأساسية The Darwin Core ، وقد طورت هذه الخطة لتقديم الحد الأدنى من وصف مجموعات التاريخ الطبيعي وقواعد بيانات الملاحظة ، كما أنها بمثابة أحد بروفايلات المعيار Z٣٩,٥٠ ، حيث طورته مجموعة المنفذين البيولوجيين للمعيار زد ٣٩,٥٠ Biological

Implementors Group (ZBIG) ، واشتهر في سياق أحد مشروعات البحوث التي تبناها متحف التاريخ الطبيعي ومركز أبحاث التنوع الإحيائي بجامعة كانساس (http://tsadev.speciesanalyst.net/) ، والمعروف بمشروع " محلل الأجناس species Analyst " .

تعرِّف الإصدارة الثانية من خطة داروين كور التي يجري تطويرها الآن بثمانية وأربعين عنصرًا من العناصر المناسبة لوصف العينات التي تتواجد ضمن المجموعات المتحفية؛ حيث تعني عدة عناصر بتعريف المكان الذي تتواجد فيه العينة بالتحديد، بما في ذلك رقمها ورمز الجهة التي تملكها أو المجموعة التي تتواجد فيها، أما بالنسبة للمعلومات التصنيفية التي تخص العينة فتسجل ضمن العناصر الخاصة بكل من "الاسم العلمي للعينة" و"الترتيب الطبقي التصنيفي الكامل taxonomic full/hierarchy" لها بدءاً من اسم المملكة[١] التي تتبعها إلى الأجناس الفرعية التي ترد تحتها ، وكذلك ضمن العناصر الخاصة بأسماء "منشئي الأسماء العلمية" ، واسم الشخص الذي طبقها على الكائن الحي أو العينة . وهناك مجموعة عناصر أخرى تعرِّف القائم على جمع العينة ، وتاريخ جمعها والوقت الذي جمعت فيه. كما أن هناك عناصر جغرافية تشتمل على ترتيب طبقي للأسماء الجغرافية (أي ذكر اسم القارة ، ثم الدولة ، ثم الولاية ، ثم المدينة) ،ومعلومات عن خطوط الطول والعرض ، والحدود ، والارتفاع ، والعمق . يضاف إلى ذلك

(١) إحدى ممالك العالم الطبيعي الثلاث [المترجم].

بعض العناصر الأخرى المتنوعة التي تشمل نمط التجهيز (مثال ذلك أن تتاح هذه العينة على شريحة) ، وعلاقة هذه العينة بالمواد الأخرى .

ويقوم نموذج " محلل الأنواع " على الجمع بين طريقة البحث المعتمدة على المعيار ٣٩,٥٠، والمجموعات المسترجعة المهيأة وفقاً للغة الترميز الموسعةXML- formatted ، حيث يمكن إدماج المعلومات الجغرافية في المجموعة الناتجة [عن البحث] في برمجية نظام المعلومات الجغرافية المعياري من أجل رسم خريطة لتوزيع الملحوظات التي سجلت عن نوع أو عينة محددة . ويعد النظام الموسوم "فش نت fishnet " من النماذج الجيدة على تقديم إمكانات البحث والاسترجاع الفيدرالي" لنموذج محلل الأنواع " (http://www.speciesanalyst.net/fishnet/)"، حيث يمكن من خلاله إجراء بحث في تسجيلات من ثلاث وعشرين مجموعة من التسجيلات الخاصة بعينات الأسماك ، ثم استرجاعها ، وإدخالها في إحدى برمجيات المعالجة في نظام المعلومات الجغرافية ، مثل برمجية ESRI's ArcView ، أو ArcMap ، وبعدها يقوم النظام بعرض خريطة للمنطقة الجغرافية تبين الأماكن التي تتواجد فيها العينات التي تم تجميعها .

READINGS :

- Federal Geographic Data Committee (home page). Available at http://www.fgdc.gov/. Accessed ٣٠ July ٢٠٠٢.

  The FGDC site contains a sidebar link to "Metadata:' a page of information about FGDC metadata, including a link to the

CSDGM document. The NBII biological profile is also available as a link from these web pages, at http://www.fgdc.gov/standards/status/subS-٢.html.

- Federal Geographic Data Committee. Content Standard for Digital Geospatial Metadata Workbook, version ٢,٠. (May ١،٢٠٠٢). Available at http://www.fgdc.gov/metadata/meta_workbook.html.

  This manual contains a complete copy of the CSDGM, annotated with additional information about the intended use of selected elements.

- Frondorf, Anne F., Matthew B. Jones, and Susan Stitt. "Linking the FGDC Geospatial Metadata Content Standard to the Biological/Ecological Sciences." In Proceedings of the Third IEEE META-DATA Conference, April ٦-٧, ١٩٩٩. Available at

  http://www.computer.org/proceedings/meta/١٩٩٩/papers/٤/afrondorf.html.

  An account of the development of the NBII Biological Metadata Standard, written before the standard was finalized as a profile of the CSDGM.

- Niemann, Brand L "Creating and Evaluating Metadata for a Digital Library of the State of the Environment." In Proceedings of the Third IEEE META-DATA Conference, April ٦-٧, ١٩٩٩. Available at http://www.computer.org/proceedings/meta/I٩٩٩/papers/٧٣lbniem

ann.htm.

How a digital library of the state of the environment was built by harvesting Dublin Core and FGDC lite metadata from web pages.

- USGS Biological Resources. An Image Map of the Content Standard for Digital Geospatial Metadata, version ٢-١٩٩٨ (FGDC-STD-OOI June ١٩٩٨). Available at http://biology.usgs.gov/fgdc.metadata/version٢/.

  A graphical representation of the CSDGM, using color coding to show which sections, compounds, and elements are mandatory, mandatory if applicable, or optional.

الفصل الخامس عشر

مبادرة توثيق البيانات

مبادرة توثيق البيانات Data Documentation Initiative هي جماعة دولية تتألف من منتجي بيانات العلوم الاجتماعية ، والأرشيفيين المعنيين بتوثيق جهود البحث العلمي في هذا المجال (http://www.icpsr.umich.edu/DDI/)، ويشير المختصر DDI إلى كل من لجنة مبادرة توثيق البيانات التي تبنت هذه المبادرة بشكل رسمي ، ومعيار ما وراء البيانات الذي طورته تلك اللجنة بهدف وصف مجموعات بيانات العلوم الاجتماعية .

وتضم مجموعات بيانات العلوم الاجتماعية كلاً من البيانات الإحصائية، ونتائج الدراسات المسحية ، والإحصاءات الصحية ، ونتائج الاستفتاءات ، وما شابه ذلك من ملفات البيانات التي قامت بتكوينها مؤسسات تجميع المعلومات المرمزة الضخمة aggregations of coded information. وتجدر الإشارة في هذا الصدد إلى أنه ثمة فوارق واضحة بين منتجي قواعد بيانات العلوم الاجتماعية من جهة ومستخدميها أو المستفيدين منها من جهة أخرى ؛ فالمنتجون غالباً ما يكونون عبارة عن هيئات حكومية ومنظمات استطلاعات الرأي ، بينما يكون المستفيدون هم الباحثون الأكاديميون ، كما أن مجموعات البيانات عادة ما تحفظ في مستودعات مركزية ضخمة كتلك التي يديرها ائتلاف الجامعات للأبحاث السياسية والاجتماعية Inter-University Consortium for Political and Social Research (ICPSR) ، ومركز روبر لأبحاث الرأي العام the Roper Center for Public Opinion Research .

ونتيجة لضعف عرى التواصل بين منتجي البيانات والمستفيدين منها ، وبحكم الطبيعة غير ذاتية التفسير non-self explanatory التي تتسم بها البيانات في هذا المجال ، يكتسب التوثيق أهمية حيوية في هذا الصدد . ولقد

كان الشكل المثالي الذي يتم من خلاله وصف مجموعات البيانات في العلوم الاجتماعية هو ما يعرف بأدلة الترميز codebooks [(١)]، حيث تشتمل على معلومات عن بنية ملفات البيانات ومحتواها ونمط إخراجها ، بما في ذلك التعريف بالمتغيرات ودلالات القيم المرمزة . ولقد تجسد الهدف الأساسي لمبادرة توثيق البيانات في محاولة استبدال الصيغة المتفق عليها والمعمول بها في أدلة الترميز الإلكترونية، والمعروفة بصيغة أوزيريس[(٢)] OSIRIS المتداولة منذ سبعينات القرن العشرين بصيغة أحدث منها . وانطلق العمل على تطوير تلك المبادرة في عام ١٩٩٥م بتمويل من ائتلاف الجامعات للأبحاث السياسية والاجتماعية ICPSR في بداية الأمر، ثم بتمويل من المؤسسة الوطنية للعلوم the National Science Foundation (NSF) بعد ذلك. وفي عام ١٩٩٩م اختبرت النسخة التجريبية beta version من هذه المبادرة،

(١) دليل الترميز : في سياق هذا النص هي وثيقة تشتمل على معلومات عن بنية ومحتوى وإخراج ملف بيانات العلوم الاجتماعية [المترجم].

(٢) ينبغي في هذا الصدد أن نفرق بين برمجية أوزوريس OSIRIS ، ودليل ترميز أوزوريس OSIRIS codebook، فأولهما برنامج إحصائي مشابه لبرنامجي SPSS و SAS تتوافر له خصائص قوية لإدارة البيانات وقد أعد هذا البرنامج معهد البحوث الاجتماعية في جامعة ميتشجن ، وقد كان ائتلاف الجامعات للأبحاث السياسية والاجتماعية -Inter University Consortium for Political and Social Research (ICPSR) في أول عهده يقوم بتوزيع العديد من الدراسات التي يعدها في صيغة أوزوريس مصحوبة بدليل ترميز خاص مقروء آليًا وملفات بيانات مقروءة آليًا كذلك عن طريق هذا البرنامج ، أما دليل ترميز أوزوريس فهو دليل ترميز مقروء آلياً مكتوب في صيغة الترقيم الثنائي binary يستخدم مع برمجية أوزوريس [المترجم] .

ثم أطلقت إصداراتها الأولى ١,٠ version في مارس من عام ٢٠٠٠م.

ومما تجدر الإشارة إليه أن هذه المبادرة نفذت كمعرف لنوع الوثيقة في صيغة لغة الترميز الموسعة XML DTD ، ووثقت عبر مكتبة تاجية tag library [١]خاصة بها. ويعرض الشكل رقم (١٥ / ١) مثالاً لتعريف أحد العناصر في المكتبة التاجية الخاصة بمبادرة توثيق البيانات .

تتكون مبادرة توثيق البيانات من خمسة أقسام رئيسية ، جاءت على النحو التالي:

- القسم الأول <docDscr> - أي وصف الوثيقة : ويشتمل على البيانات الببليوجرافية لوثيقة مبادرة توثيق البيانات .
- القسم الثاني <stdyDscr> - أي وصف الدراسة : ويشتمل على معلومات عن عملية جمع البيانات ، أو الدراسة ، أو التجميع الذي يتم وصفه عن طريق دليل الترميز الخاص بمبادرة توثيق البيانات .
- القسم الثالث <fileDscr> - أي وصف ملفات البيانات: ويعنى بوصف سمات ومحتويات ملفات البيانات التي تكوّن الدراسة .
- القسم الرابع <dataDscr> - أي وصف المتغير : ويعنى بوصف كل متغير من المتغيرات تضمنها ملف البيانات .

(١) مكتبة أو دليل التيجان أو العناصر: وثيقة تتضمن قائمة بأسماء عناصر البيانات وخصائصها مرتبة ترتيباً هجائياً ، مع تعريفاتها وقواعد استخدامها في لغة الترميز الموسعة واللغة العامة الموحدة للترميز [المترجم].

- القسم الخامس <fileDscr> - أي وصف المواد الأخرى ذات الصلة بالدراسة : ويعنى بتحديد التقارير والمطبوعات والمواد الأخرى التي لها علاقة معينة بالدراسة.

ومن الملاحظ أن القسم الأول - قسم وصف الوثيقة - يصف وثيقة مبادرة توثيق البيانات بشكل عام ، ويتكون من قسمين رئيسيين : أولهما قسم الإرجاعة الببليوجرافية أو الاستشهاد <citation> ، ويعني بوصف دليل الترميز الإلكتروني، وثانيهما قسم الوثيقة المصدرية <docScr> ويعني بوصف الوثيقة المصدرية التي قد تكون عبارة عن دليل ترميز مطبوع، أو دليل ترميز إلكتروني موجود في صيغة أخرى ، وذلك أشبه بما يحدث في ترويسة مبادرة ترميز النص TEI header التي تشتمل على عنصر الوصف المصدري <sourceDesc> الذي يعنى بوصف النص المصدري داخل عنصر وصف الملف <fileDesc> المعني بوصف النص الإلكتروني ،غير أن هذين العنصرين - أي عنصر الارجاعة الببليوجرافية وعنصر الوثيقة المصدرية - يأتيان في سياق مبادرة توثيق البيانات في المستوى نفسه بداخل قسم وصف الوثيقة <docDscr> [١] .

ويتضمن عنصر الإرجاعة الببليوجرافية <citation> على بيانات العنوان، وبيان المسؤولية، وبيانات الإنتاج والتوزيع، وبيانات السلسلة، وبيانات الإصدارة، بالإضافة إلى إرجاعة ببليوجرافية تحيل إلى وثيقة

(١) لاحظ الفرق الدقيق بين التاجين ؛ إذ يكون تاج القسم الأكبر أي وصف الوثيقة هكذا <docDscr> ، بينما يكون تاج وصف الوثيقة المصدرية هكذا <docScr> أي بزيادة حرف D في التاج الأول [المترجم].

مبادرة توثيق البيانات، ومعلومات الاقتناء (أي توضيح المكان الذي توجد به وثيقة مبادرة توثيق البيانات)، هذا بالإضافة إلى قسم خاص بالتبصرات تسجل فيه أية معلومات إضافية عن الإرجاعة البيليوجرافية. مع ملاحظة أن بيانات العنوان التي تتواجد بداخل هذه الإرجاعة البيليوجرافية يمكن أن يسجل بها العنوان الرئيسي <titl>،

Variable Label

<labl> 4.2.2 (Generic element A.2)

Description: A descriptive phrase which defines the variable. The length of this phrase may depend on the statistical analysis system used (e.g., some version of SAS permit 40-character labels while some versions of SPSS permit 120 characters. A "level" attribute is included to permit coding of the level to which label applies, i.e., the study level, the file level (if different from study), the record group, the variable group, or the variable level. Vendor attribute provided to allow for specification of different labels for use with different vendors' software.

Remarks: Whenever possible this element should be used instead of 4.2.15 (Variable Text, 'txt') in order to facilitate the creation of statistical analysis software labels.

Example:

<var><labl>Why No Holiday-No Money</labl></var>

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, level, vendor
- Contains: #PCDATA, Link to other element(s) within the codebook.

ال ... وثيق

ID

ber ... وثيق البيانات تحرص بشكل صريح على ربط عناصر البيانات الفردية داخل عنصر

الإرجاعة الببليوجرافية بعناصر دبلن كور ، كما توصي كذلك بتوفير القيم الخاصة بتلك العناصر .

وكما أشرنا من قبل ، يصف عنصر الوثيقة المصدرية <docScr> وثيقة مبادرة توثيق البيانات ، ومن ثم فإنه يشتمل على العناصر الفرعية نفسها التي يتضمنها عنصر الإرجاعة الببليوجرافية ، ولكن بإعادة توضيحها لتشير إلى أنها خاصة بالوثيقة المصدرية ، وهناك عنصران فرعيان من العناصر الفرعية عالية الرتبة high order* يتواجدان بداخل قسم وصف الوثيقة هما : عنصر الدليل <guide> ، وعنصر حالة الوثيقة <docStatus>، وأول هذين العنصرين عبارة عن قائمة نصية بالمصطلحات المعتمدة في وثيقة مبادرة توثيق البيانات، مع تعريفاتها، بهدف مساعدة المستفيد على فهم هذه الوثيقة، أما ثانيهما - أي حالة الوثيقة - فيصف بشكل نصي كذلك الحالة التي أنتجت فيها وثيقة مبادرة توثيق البيانات، كأن تكون هذه الوثيقة وثيقة تمهيدية أو غير مكتملة .

ويشتمل القسم الرئيسي الثاني - قسم وصف الدراسة - على المعلومات الخاصة بالدراسة التي أثمرت مجموعة بيانات العلوم الاجتماعية التي يتم وصفها ، ولهذا القسم ستة أقسام فرعية رئيسية هي : الإرجاعة الببليوجرافية <citation> ، ومعلومات الدراسة <stdyInfo>، ومنهجية إعدادها <method>، وطريقة إتاحتها أو الوصول إليها <dataAcc> ، والمواد الأخرى ذات الصلة <otherStdyMat>، والتبصرات <notes>. ويلاحظ في هذا الصدد أن القسم الفرعي " الإرجاعة الببليوجرافية " هو العنصر المغلف نفسه

wrapper [١] مع مجموعة العناصر الفرعية التي حددت في قسم وصف الوثيقة، إلا أن هذين العنصرين في هذا الموضع يختصان بوصف الدراسة، وليس بوصف دليل ترميز مبادرة توثيق البيانات ، حيث تشير قيمة عنصر المؤلف - مثلاً - هنا إلى الباحث الرئيسي للدراسة investigator ، وليس إلى منشئ دليل الترميز . كذلك هناك القسم الفرعي الخاص بمجال الدراسة (<stdyInfo>) الذي يشتمل على المعلومات التي تصف المحتوى الفكري للدراسة، ومجال تغطيتها الجغرافية والزمنية . مع ملاحظة أن رؤوس الموضوعات يمكن أن تسجل ككلمات مفتاحية في عنصر " الكلمة المفتاحية <keyword> "، أو كمصطلحات يعتمد في صياغتها على إحدى اللغات المقيدة ومن ثم تسجل في العنصر الموسوم " الفئة الموضوعية <topcClas>" ، الذي يسجل فيه اسم هذه اللغة في خاصية معنية بذلك ، وهي الخاصية المسماة " اللغة المقيدة vocab " ، أما المحدد الموحد الخاص لتلك اللغة المقيدة فيسجل في الخاصية المعنية بذلك، وهي الخاصية المسماة " المحدد الموحد للغة المقيدة "vocabURI".

أما القسم الفرعي " المنهجية والمعالجة <method> في قسم وصف الدراسة فتُعرّف في سياقه العناصر التي تقدم معلومات عن المنهجية المتبعة في إعداد الدراسة بما في ذلك المعلومات الدقيقة عن طريقة جمع البيانات، مثل أسلوب جمع البيانات ، وطبيعة أداة البحث، وإجراءات تحديد عينة الدراسة إلى غير ذلك من بيانات مما يدخل في هذا الجانب ، كما يشتمل هذا القسم الفرعي كذلك على المعلومات التقييمية أو التقديرية appraisal

(١) يقصد بالعنصر المغلّف ، وأحياناً يسمى الحاوي container ، عنصر بيانات كبيرًا ترد تحته عناصر بيانات فرعية تابعة له أو لأحد فروعه [المترجم].

الخاصة بالدراسة، مثل معدل الاستجابة ، واحتمالات الخطأ في العينة، وحالة الدراسة . أما القسم الفرعي "كيفية إتاحة البيانات أو الوصول إليها" في قسم وصف الدراسة، فيصف الموقع الذي تتواجد فيه مجموعة البيانات ، ومدى توافرها . وأما القسم الفرعي " المواد الدراسية الأخرى" فيستخدم للتعريف بالمواد ذات الصلة التي لها علاقة ما بالدراسة التي يتم وصفها ، كالكتب والدراسات الأخرى .

ويأتي القسم الرئيسي الثالث - وصف ملف البيانات - كقسم متكرر يخص كل ملف من الملفات التي تتكون منها مجموعة البيانات . ويندرج تحت هذا القسم الرئيسي قسمان فرعيان هما : وصف الملف < fileTxt >، والتبصرات <notes>. حيث يركز أولهما على الوصف المادي لملف البيانات المادي وصفاً تفصيلياً، عن طريق تسجيل اسم الملف، ومحتواه ، وبنيته، وأبعاده، وشكل بياناته، ومعلومات تجهيزه، بالإضافة إلى معلومات عن البيانات المفقودة من هذا الملف، بينما تسجل في ثانيهما أية معلومات إضافية عن الملف ولم تكن قد وردت في سياق القسم الأول .

أما القسم الرئيسي الرابع - قسم وصف المتغيرات - فيستخدم لوصف كل متغير من المتغيرات التي يتضمنها ملف البيانات ، ويندرج تحت هذا القسم ثلاثة أقسام فرعية هي:مجموعة المتغيرات <varGrp>، والمتغير <var>، والتبصرات <notes> ؛ حيث يجمع أولها المتغيرات التي يتقاسمها عامل مشترك، مثال ذلك الصيغ المختلفة التي يأخذها السؤال نفسه، أو الأسئلة المختلفة التي تدور حول موضوع واحد . أما عنصر " المتغير " فيصف جميع السمات المتصلة بمتغير واحد في ملف البيانات، ويشمل ذلك تقديم المعلومات

عن اسم هذا المتغير، وموضعه في ملف البيانات (بمعنى هل يوجد في بداية الملف أم في نهايته)، وماذا يصف هذا المتغير ، وهل له وزن معين، وما نوع بياناته وصيغها، ومتوسط قيمه الصالحة وغير الصالحة، وعندما يمثل المتغير إجابة أحد الأسئلة تكون هناك عناصر لتسجيل السؤال وما يتصل به من معلومات، مثال ذلك التعليمات التي يقدمها القائم بإجراء المقابلة .

ويستخدم القسم الرئيسي الأخير - قسم المواد الأخرى ذات الصلة بالدراسة - لتسجيل المواد الأخرى التي اعتمد عليها في إعداد الدراسة أو التي أفادت منها الدراسة، وخاصة في تحليلاتها، أو يستخدم لعمل روابط لتلك المواد، من هذه المواد الاستبانات المستخدمة في المسح ، وإرشادات الترميز ، وأدلة الاستخدام، بل قد يتسع مجالها لتشمل البرامج الآلية المستخدمة . وفي هذا الصدد يمكن تسجيل المواد النصية مباشرة في العنصر الفرعي المعني بذلك والمسمى "نص <txt> "، أما جداول البيانات فيتم تسجيلها في العنصر الفرعي المسمى "جدول <table> "، كما يمكن كذلك الإشارة إلى تلك المواد عن طريق إعطاء المحددات الموحدة لها URI . ويعرض الشكل رقم (١٥ / ٢) بداية عرض متاح على الإنترنت لأحد أدلة الترميز التي تم ترميزها اعتماداً على مبادرة توثيق البيانات .

وتجدر الإشارة إلى أن الإصدارة الأولى من معرف نوع الوثيقة الخاص بمبادرة توثيق البيانات DDI DTD لا تشير إلى أية عناصر إجبارية باستثناء عنصر العنوان ، غير أن موقع هذه المبادرة على الإنترنت توجد به

صـفحة تتـضمن عـدة عناصـر مفـضلة[١] recommended ، كمـا تتـضمن توصية مفادها أن يقوم كل مجتمع من المستفيدين بتطوير مجموعة عناصر مفضلة خاصة به . ومع أن مبادرة توثيق البيانات لم تحدد أية قواعد خاصة بالمحتوى ، فقد أشير فـي سياق مواصفة هذه المبادرة إلـى بعض القوائم الاسـتنادية التـي يمكـن أن يعتمد علـيها فـي صـياغة قـيم خـصائص بعـض العناصـر ، ولكن يمكن القول بشكل عام إنه ليست هناك لغات مقننة أو قوائم اسـتنادية إجبارية فـي هـذا الـصدد ، وإن كـان هـناك تـوجه مـستقبلي يـشير إلــى ضرورة تطوير لغات مقيدة " لأكبر عدد ممكن من الخصائص".

كمـا أن هـناك توجـيهات أخـرى لمـبادرة توثـيق البـيانات لدراسـة مـدى إمكانـية اسـتخدام لغـة الترميـز الموسـعة ، وإطـار وصـف المـصدرRDF ، وإمكانـية تـضمينها معايير مـا وراء البيانات الجغرافية والمكانـية ، وإضـافة نقـاط الـتحويل البينـي crosswalks إلـى خطـط ببليوجرافية أخـرى مثل خطتي جيلس GILS ومارك .

ويعد ائتلاف الجامعات للأبحاث السياسية والاجتماعية ICPSR الذي يعد أحد الأرشيفات الضخمة لبيانات العلوم الاجتماعية داعماً قوياً لصيغة مبادرة توثيق البيانات ، حيث قـام بالفعل بتــرميــز جميع جهــود الوصف الخاصة بجميع

CBS News Monthly Poll #2, August 1992 (ICPSR 6094)
August National Poll II, Republican National Convention)

View: Part 1: Document Description
Part 2: Study Description
Part 3: Data Files Description
Part 4: Variable Description
[illegible]

Document Description	
Citation	
Title:	CBS News Monthly Poll #2, August 1992
Alternative Title:	August National Poll II, Republican National Convention
Identification Number:	6094
Authoring Entity:	CBS News
Producer:	Inter-university Consortium for Political and Social Research
Copyright:	Copyright ICPSR 2000
Date of Production:	May 10, 2000
Software used in Production:	SoftQuad XMetaL

(١) يقص... ـر الإجبارية
[المتر...
أساسيات...

الشكل رقم (١٥ / ٢) الصفحة الأولى من شاشة العرض الخاصة بأحد أدلة الترميز التي تم ترميزها اعتماداً على مبادرة توثيق البيانات .

المصدر : موقع مبادرة توثيق البيانات على الإنترنت :

Data Documentation Initiative Website . "Marked up Codebooks" available at http://www.icpsr.umich.edu:٨٠٨٠/DDI/SAMPLES/٠٦٠٨٤.xml.

الدراسات التي يحصرها ووردت ضمن الفهرس الخاص به، وذلك اعتماداً على مواصفة هذه المبادرة ، كما يتضمن موقع هذه المبادرة قائمة بجميع الأرشيفات ومشروعات وصف البيانات التي تبنتها ، ولعل من أكبر المشروعات في هذا الصدد مشروع نيستار NESSTAR (أو أدوات ومصادر العلوم الاجتماعية الشبكية Networked Social Sciences Tools and Resources) - أحد مشروعات التطوير المشتركة بين كل من الخدمات النرويجية لبيانات العلوم الاجتماعية Norwegian Social Science Data Services ، وأرشيف بيانات المملكة المتحدة the U.K Data archive ، وأرشيف البيانات الدانماركي the Danish Data Archive.

ويقدم نموذج نيستار خدمات بحث واسترجاع فيدرالية عبر مستودعات بيانات موزعة متخصصة في العلوم الاجتماعية ، مع توفير أدوات متكاملة لتصفح ما وراء البيانات الخاصة بمجموعات البيانات ، ولتحليل وعرض البيانات ، فضلاً عن إمكانية تحميل مجموعات فرعية منها البيانات . وما كان من الممكن تقديم كل تلك الخدمات لولا توافر ما وراء بيانات ذات بنية فائقة تمكّن من استكشاف المصادر وعرضها وتحليل بياناتها .

ولقد كشف مشروع نيستار حقيقة مفادها أن مبادرة توثيق البيانات قد تجاوزت حدود ما كانت تطمح إلى تحقيقه المتمثل في محاولتها أن تحل محل صيغة أوزوريس المتقادمة ، كما فتح تطوير صيغة مبادرة توثيق البيانات المعتمدة على لغة الترميز الموسعة آفاقاً جديدة لعمليات التشغيل البيني التي تتراوح من مجرد البحث الفيدرالي إلى تمكين الباحث من استخدام عدة مجموعات بيانات دفعة واحدة . وفي الوقت نفسه أثار استخدام مبادرة توثيق البيانات الحاجة إلى إحداث تغييرات وتوسعات على الصيغة ذاتها، وبما يوحي بحاجتها إلى الصيانة المستمرة .

ومع كل هذه التطلعات فإن مبادرة توثيق البيانات تسعى الآن سعياً حثيثاً لتأمين مستوى مناسب من التمويل المستمر، وقد يدعوها ذلك لإحداث قدر كبير من التغييرات على بنيتها التنظيمية استجابة لذلك . وأخيراً يمكن القول إن مبادرة توثيق البيانات باتت بالفعل جزءاً من نسيج البحث في مجال العلوم الاجتماعية، كما بات مستقبلها كخطة ما وراء بيانات مأموناً .

READINGS :

- Data Documentation Initiative (home page). Available at http://www.icpsr.umich.edu/DDI/. Accessed ٣١ July ٢٠٠٢.

  The DDI website has links to the DDI DTD and tag library, background information on the DDI, and sample DDI-encoded codebooks.

- Ryssevik, Jostein, and Simon Mugrave. "The Social Science Dream Machine: Resource Discovery, Analysis, and Delivery on the Web." Paper given at the IASSIST Conference, Toronto, May ١٩٩٩. Available at http://www.nesstar.org/papers/iassist_٠٥٩٩.html.

  A discussion of the NESSTAR vision, with some reference to the use of DDI.

الفصل السادس عشر

ما وراء البيانات الإدارية

في الوقت الذي تهدف فيه ما وراء البيانات الوصفية إلى المساعدة في البحث عن مصدر معين من مصادر المعلومات واستكشافه والتعرف إليه ، تسعى ما وراء البيانات الإدارية إلى تيسير سبل إدارة هذا المصدر . وتشمل وظائف الإدارة في أحسن أحوالها تتبع المصدر عبر مختلف مراحل معالجته، وضبط إتاحته أو الوصول إليه ، والتحقق من المسؤوليات المتصلة به، ومنح تصاريح استخدامه أو الإفادة منه . ومع أن هذا الأمر قد يبدو مبسطاً أكثر من اللازم ، فيمكن القول إن ما وراء البيانات الوصفية توجه أساساً لخدمة المستفيد الفعلي أو المحتمل من المصدر ، بينما توجه ما وراء البيانات الإدارية لخدمة صاحب هذا المصدر أو القائم على أمره .

ومن الواضح أنه ليس هنالك فارق قاطع بين ما وراء البيانات الوصفية وما وراء البيانات الإدارية ؛ ذلك لأن معظم عناصر البيانات التي تعرّف في سياق هاتين الفئتين يمكن استخدامها في سياق وصفي أو في سياق إداري ، وتبعاً لذلك فليس هنالك فارق قاطع بين خطط ما وراء البيانات الوصفية وخطط ما وراء البيانات الإدارية ؛ ذلك لأن معظم الخطط الوصفية - التي أشرنا إليها في الفصول السابقة - تحتوي إلى حد ما على عناصر وظيفتها إدارية في المقام الأول، أضف إلى ذلك أن كثيراً من الخطط عادة ما تضع في الاعتبار عند تصميمها أن توجه في الوقت نفسه لخدمة أغراض وصفية وإدارية ، ولعل من أفضل الأمثلة على ذلك محدد مواقع المعلومات الحكومية المعروف بجيلس الذي روعي في تصميمه أن يكون ذا هدف مزدوج ؛ حيث يسعى من ناحية إلى تيسير سبل الوصول إلى المعلومات الحكومية من ناحية ، وإلى مساندة إدارة المحفوظات أو الأرشيفات الجارية

من ناحية أخرى ، ومع ذلك يلاحظ أن الغالبية العظمى من الخطط تتجه إلى التركيز على أحد هذين البعدين ، وبالشكل الذي يدفعنا إلى تصنيفها إلى خطط ما وراء بيانات وصفية ، وخطط ما وراء بيانات إدارية .

ومن الملاحظ أن جوانب ما وراء البيانات الإدارية قد تحظى بالتركيز من جانب خطط معينة وبشكل قد يختلف عن التركيز الذي تحظى به في خطط أخرى ؛ حيث نجد أن خطط ما وراء البيانات الخاصة بالحفظ preservation تركز على عناصر تعنى بضمان حفظ مصادر المعلومات حفظاً طويل الأجل، وبمدى قابلية استخدام تلك المصادر أو الإفادة منها، بينما تركز ما وراء البيانات الفنية على وصف عملية إنشاء الكيانات الرقمية، ووصف خصائصها المادية، وغالباً ما تكون هذه الفئة من ما وراء البيانات - أي الفنية - مكوناً أساسياً من وراء بيانات الحفظ، وهناك كذلك ما وراء البيانات الخاصة بالحقوق rights metadata - التي تعني بتوثيق وإدارة الحقوق وتصاريح الاستخدام المرتبطة بمصادر المعلومات . ونلقي الضوء في هذا الفصل على الخطط العامة للفئتين الأوليين - أي ما وراء البيانات الإدارية، وما وراء بيانات الحفظ - ونرجئ الحديث عن الفئة الثالثة - أي ما وراء بيانات الحقوق - إلى الفصل الأخير من هذا الكتاب (الفصل الثامن عشر) .

الخطط الإدارية :

على عكس ما عليه الحال بالنسبة للخطط الوصفية التي تتسم بكثرتها، ليست هنالك خطط ما وراء بيانات إدارية عامة تحققت لها صفة المعايير الرسمية أو المجتمعية؛ وقد يرجع ذلك إلى أن الحاجة إلى التشغيل البيني

interoperability بين التطبيقات الإداريـة لم تبلغ درجة الإلحاح التي وصلت إليها الحاجـة إلـى التشغيل البينـي لأغراض البحث أو البحث البيني search interoperability ؛ فالجهـات المعنـية بـصيانة مـصادر المعلـومات عـادة مـا يكـون لـديها أنظمـتها الداخلية الخاصـة بـإدارة مصادرها عن طريق ما وراء بيانات أعدت خصيصاً لهذه التطبيقات أو لهذا الغرض application-specific metadata schemes، كمـا أنـها نادراً ما تحتاج إلى تقاسم ما وراء البيانات هذه أو تـبادلـها مـع جهـات أخـرى، بل قد لا يكون لديها ما يدفعها لإجراء عمليات بحـث لمستودعات مـا وراء بيانات إداريـة تديرها جهات أخرى . ومع ذلك فـثمة جهـود عديـدة نـشطة تـسعى لتطويـر مجمـوعات مقنـنة أو قياسـية مـن عناصر ما وراء البيانات الإدارية سنشير على اثنتين منها في الفقرات التالية .

تتجسد أول الجهود البارزة في هذا الصدد فيما تقوم به مبادرة دبلن كور لمـا وراء البيانات مـن تخصيصـها مـا وراء بيانات إداريـة مـرتبطـة ارتـباطًا وثـيقًا بمـا وراء البـيانات الوصـفية، والتـي تعد فـي حد ذاتهـا بمـثابة مـصدر معلومات يحتاج إلى إدارة (غالباً ما يطلق على مثل تلك البيانات التي تصف مـا وراء البيانات تسمية غير مستساغة هي met-metadata. فهناك خطة دبلن كور الإدارية Administrative Dublin Core - أو ما تعرف اختصاراً A-Core ، وهـي مسودة خطـة مبكـرة جداً كانـت قد طـورتها مجموعة عمل ما وراء البيانات الإدارية التابعة لمبادرة دبلن كور ، ويعرّف التصور المقترح DC-٩ ـ الذي نـوقش فـي إحدى ورش العمل التي خصصت لذلك وكانت بعنوان " DC-٩" والتـي صـاحبت المؤتمـر التاسـع لـدبلن كـور وتطبـيقات مـا وراء

البيانات المنعقد في طوكيو (DC-٢٠٠١: International Conference on Dublin Core and Metadata Applications) يعرّف خطة تتألف من مجموعة عناصر ما وراء بيانات تقع في ثلاث فئات، تسجل أولاها الأعمال الإجرائية Action التي أجريت على أي من تسجيلات ما وراء البيانات الوصفية (مثل إنشاء التسجيلة، أو تعديلها، ومراجعتها)، ومن أمثلة ما يتم تسجيله في هذا الصدد : نوع الإجراء، وتاريخه، والمسؤولون عنه، ومعلومات عن الجهات التي ينتسبون إليها، ووسائل الاتصال بهم . أما الفئة الثانية من العناصر فتعني بتسجيل المعلومات الأساسية أو الثابتة nontransactional الخاصة بتسجيلة ما وراء البيانات، مثل : لغتها، وحالتها، وموقعها، وملكيتها . ثم تأتي الفئة الثالثة من العناصر، وهي خاصة ببعض العناصر التي لا تتصل بتسجيلة ما وراء بيانات محددة، وإنما بمجموعة أو دفعة batches من التسجيلات التي يتم تبادلها كوحدة واحدة، وتشمل هذه العناصر رمزاً يعرف بقاعدة البيانات المستهدفة التي سترسل إليها دفعة التسجيلات، واسم أو رمز الجهة المرسلة، واسم الملف، والصيغة الفنية للدفعة (مثل لغة الترميز الموسعة، أو لغة ترميز النص الفائق)، وصيغتها الببليوجرافية (مثل دبلن كور، أو مارك ٢١، أو غيرهما)، ومجموعة حروفها، أما آخر ما يسجل من عناصر فهو ما يحدد طبيعة الإجراء المتخذ بشأن تلك الدفعة (مثال ذلك : التعليق، أو التعديل، أو الحذف).

لقد قصد بدبلن كور الإدارية أن تطبق في إدارة جميع أنواع ما وراء البيانات، وليس ما وراء بيانات دبلن كور وحدها، كما أنها - وكما هو الحال مع دبلن كور نفسها - تعنى بتعريف الجوانب الدلالية semantics ، كما يمكن

تنفيذها في أي تركيبة، ولكن نظراً لأن هناك مجموعة معينة من العناصر يجب أن تجمع مع بعضها بعضاً حتى تكون ذات مغذى (مثل نوع العمل ، وتاريخه ، والطرف المسؤول عن أدائه)، فقد رؤي أن يكون من المناسب اختيار لغة الترميز الموسعة حيث تسمح بتجميع تلك العناصر على نحو واضح، وهي بذلك تفضل تركيبة لغة ترميز النص الفائق التي لا تتيح تلك الإمكانية.

أما الشكل الثاني من الجهود فيتمثل في إحدى المجموعات التي تعني بتخصيص مجموعة عامة من عناصر البيانات الإدارية ، وهذه المجموعة - التي لا تزال في مراحلها الأولى - معنية بالجوانب الخاصة بإدارة تراخيص استخدام المصادر الإلكترونية المتاحة تجارياً ، وإدارة إتاحتها أو الوصول إليها . ويقود هذه المجموعة كل من آدم تشاندلر Adam Chandler من جامعة كونريل ، وتم جويل Tim Jewell من جامعة واشنطن، مع ملاحظة أن هذه المجموعة ليس لها اسم رسمي، ولكن تدير قائمة نقاش عبر موقع لها على الإنترنت عنوانه :

" نقطة ويب ارتكازية لتطوير ما وراء بيانات إدارية لإدارة المصادر الإلكترونية

A Web Hub for Developing Administrative Metadata for Electronic Resources Management".

(http://www.library.cornell.edu/cts/elicensestudy/home.html)

ويشير المشاركون في هذه المجموعة إلى أنه لعدم قدرة الجيل الحالي

من النظم المتكاملة للمكتبات على توفير نظم فرعية خاصة بإدارة المحتوى الإلكتروني المرخص به ، تقوم كثير من المكتبات بتطوير نظم قواعد بيانات خاصة بها لهذا الغرض ، وتتمثل الوظائف التي تقوم بها هذه النظم في تسجيل قرارات اختيار مصادر المعلومات الإلكترونية ، ومتابعة إجراءات الاشتراك فيها ، أو تجديد الاشتراكات القديمة ، وتوثيق المصادر الفردية التي تتضمنها المنتجات المجمعة، وتوثيق ضوابط إتاحتها وقيود استخدامها ، وتسجيل فترات عدم إتاحتها، إلى غير ذلك من القضايا المتصلة بتقديم الخدمات المتاحة على الخط المباشر.

من هنا جاء تركيز جهود ما وراء البيانات على محاولة تقنين standardizing المصطلحات المتصلة بضوابط إتاحة المصادر الإلكترونية وتراخيص استخدامها، ورأت أنه من قبيل المسلمات أن يتوافر معجم بيانات معياري standard data dictionary يضبط المصطلحات المتداولة في هذا الصدد ، وهذا من شأنه الارتقاء بعرى التواصل بين الأقسام المعنية بإدارة المصادر الإلكترونية بالمكتبات ، وتمكينها من تبادل المعلومات عن ضوابط عمليات الترخيص باستخدامها وإجراء الاتفاقيات فيما بينها في هذا الصدد ، إضافة إلى أن ذلك من شأنه أن يشجع موردي النظم المتكاملة على تطوير نظم تراعي الاحتياجات الفعلية المكتبات .

ما وراء البيانات الفنية :

نظراً لأن ما وراء البيانات الفنية تعني بتوثيق عملية إنشاء الملفات الرقمية ووصف خصائصها ، فإن طبيعة المعلومات المتصلة بذلك تتفاوت

تفاوتاً واضحاً حيث إنها ترتبط بنوع الملف الذي سيوصف ، ومن ثم فإن خطط ما وراء البيانات الفنية تميل إلى أن تكون خططًا خاصة بنوع أو بصيغة الملف format-specific .

وتعد خطة " ما وراء البيانات الفنية للصور الرقمية الثابتة Technical Metadata for Digital Still Images "(١) إحدى الخطط التي في طريقها إلى الحصول على صفة المعايير الرسمية من قبل كل من المنظمة الوطنية لمعايير المعلومات NISO ، والجمعية الدولية لإدارة المعلومات والصور AIIM International . ومن الواضح أن مواصفة هذه الخطة - التي كانت بمثابة مسودة معيار تجريبي حتى ديسمبر من عام ٢٠٠٣م - تتعامل مع صيغ الصور غير المتحركة، مثل صيغ : تيف TIFF ، وجي بيج JPEG ، وجيف GIF ، وبي دي إف PDF . وتحدد هذه المواصفة بالنسبة لكل عنصر من عناصر البيانات التي تعرفها ما يلي : نوع البيانات، وهل هذا العنصر من العناصر الإجبارية أم من العناصر الإجبارية عندما يكون ذلك ضرورياً، أم من العناصر المفضلة، أم من العناصر الاختيارية، وهل هو من العناصر القابلة للتكرار، وهل هناك قيود على استخدامه ، بمعنى هل يستخدم من قبل النظام بشكل عام، أم من قبل مدير النظام، أم من قبل المستفيد النهائي، كما تشير المواصفة إلى القوائم الاستنادية التي ينبغي الاعتماد عليها عندما

(١) "Data Dictionary-Technical Metadata for Digital Still Images," NISO Z٣٩.٨٧-٢٠٠٢ AIIM ٢٠-٢٠٠٢, available at

http://www.niso.org/standards/resources/Z٣٩_٨٧_trial_use_pdf. Accessed ٥ July ٢٠٠٢

يتطلب الأمر أن يعتمد في صياغة نوع البيانات على إحدى القوائم الاستنادية للقيم المرمزة أو القيم النصية ، وفي هذا الصدد تتضمن المواصفة كثيرًا من إرشادات الاستخدام، فضلاً عن الإحالة إلى عناصر البيانات المشابهة والمعرفة في المواصفات الأخرى ذات الصلة بهذه المواصفة، مثل مواصفة صيغة تيف .

وقد قسمت عناصر ما وراء البيانات في سياق هذه المواصفة في أربع مجموعات جاءت على النحو التالي : مجموعة المحددات parameters ، ومجموعة إنشاء الصور ، ومجموعة تقييم الأداء، ومجموعة تاريخ التغيير . وتشتمل مجموعة المحددات الأساسية على المعلومات الخاصة بالصيغة، مثل : نوع ملف المايم[١] MIME ، ومعدل ضغطه، ومعلومات الملف، مثل : حجمه، ورقم التدقيق checksum الخاص به . أما مجموعة العناصر الخاصة بإنشاء الصورة فتتضمن المعلومات التفصيلية عن كيفية التقاط الصورة، وهل تم ذلك عن طريق الكاميرا الرقمية ، أو عن طريق المسح الضوئي للمصدر التناظري[٢] . أما عناصر مجموعة تقييم الأداء فقد اختيرت لتكون بمثابة " مقاييس لتقدير درجة دقة المخرجات (عند الاستخدام الحالي)، وتقدير درجة دقة أساليب الحفظ ، وبخاصة أساليب نقل أو تحويل البيانات Migration (عند الاستخدام المستقبلي) . ومن ثم فهي تضم مقاييس الحيز spatial metrics، مثل : طول الصورة وعرضها،

---

(١) MIME type : اسم رسمي لأنواع وسائط الإنترنت : مجموعة معيارية للمصطلحات المستخدمة لتعريف الصيغ الرقمية [المترجم].

(١) تقصد المصدر الأصلي غير الرقمي [المترجم].

والمقاييس غير المكانية، مثل : عدد مكونات اللون لكل بكسل . ثم تأتي بعد ذلك مجموعة عناصر تاريخ التغيير التي اختيرت لتسجيل جميع عمليات التحرير التي أجريت على الصورة، بما في ذلك بيان المسؤول عن هذه العمليات ، وتاريخ ووقت إجرائها، والبرمجية المستخدمة في هذا الصدد .

ولقد قامت مكتبة الكونجرس بإعداد خطة بصيغة لغة الترميز الموسعة خاصة بالجوانب الدلالية التي أقرتها المنظمة الوطنية لمعايير المعلومات والجمعية الدولية لإدارة المعلومات والصور ، وتعرف هذه الخطة اختصاراً بنيزوميكس NISO MIX أو خطة المنظمة الوطنية لمعايير المعلومات لما وراء بيانات الصور في صيغة لغة الترميز الموسعة NISO Metadata for Image in XML Schema . [١] وتجدر الإشارة هنا إلى أن هذه الخطة والأدوات ذات العلاقة المتصلة بتنفيذ معجم البيانات متاحة على موقع مكتب التطوير الشبكي ومعايير مارك على الوب (http://www.loc.gov/standards/mix/) .

وفي السياق نفسه تبنى المشروع الأولي لمكتبة الكونجرس لخطط الصوتيات والمرئيات الرقمية the Library of Congress's Digital Audio-Visual Preservation Prototype project جهودًا مكثفة لتعريف مجموعة عناصر ما وراء بيانات فنية للصوتيات والمرئيات الرقمية

(٢) MIX Metadata for Image in XML Schema,
خطة ما وراء بيانات الصورة بصيغة لغة الترميز الموسعة : عبارة عن خطة في صيغة لغة الترميز الموسعة لتمثيل الدلالات الخاصة بمعيار المنظمة الوطنية لمعايير المعلومات و الجمعية الدولية لإدارة المعلومات والصور لما وراء البيانات الفنية للصور الرقمية الثابتة [المترجم].

(http://www.loc.gov/rr/mopic/avprot/avprhome.html) ؛ فهـناك خطـة أوديومـيد AUDIOMD Audio Technical Metadata Extension Schema ، التـي تشتمل على سبعة وثلاثين عنصراً من العناصر عالية المستوى تعنى بوصف ملفات الصوت الرقمية ، ومصادرها التناظرية أو الرقمية عندما يكون وصفها ضرورياً ؛ وهناك كذلك خطـة فيدويوميد VEDIOMD Vedio Technical Metadata Extension Schema - التـي تشتمل علـى ستة وثلاثين عنصراً من العناصر عالية المستوى تعني بوصف ملفات الفيديو الرقمية ومصادرها التناظرية أو الرقمية.

ومن الملاحظ أن موقع هذا المشروع على الوب يثير الاهتمام لا لكثرة ما يتضمنه من روابط لخطط ما وراء بيانات فنية فحسب، وإنما لوجود روابط إلـى أحد جداول التطبيقات الذي يبين الكيفية التـي نفذت بها تلك العناصر في إحدى قواعد البيانات العلاقية التـي أعدت بالفعل لدعم إدارة وظائف الحفظ .

وكما هو الحال مع خطـة ما وراء بيانات الصور فـي صيغة لغة الترميز الموسعة الخاصة بما وراء البيانات الفنية للصور الرقمية الثابتة التي تبنتها كل من المنظمة الوطنية لمعايير المعلومات والجمعية الدولية لإدارة المعلومات والصور (انظر الفصل ١٧)؛ فقد عرفت خطتا أوديوميد وفيديوميد على أنهما خطتا توسعة فـي صيغة لغة الترميز الموسعة لمعيار ترميز ونقل ما وراء البيانات(١) XML extension schema for METS .

(١) MITS Metadata Encoding and Transmission Standard [المترجم] .
معيار ترميز ونقل ما وراء البيانات: مواصفة لما وراء البيانات البنائية أو الهيكلية [المترجم].

ما وراء بيانات الحفظ :

أدى الاهتمام المتزايد بأرشفة المواد الرقمية وحفظها حفظاً طويل الأجل إلى التفكير الجدي في تطوير خطط ما وراء بيانات تساعد على إدارة عمليات الحفظ ؛ ففي عام ١٩٩٨م أصدرت مجموعة العمل الخاصة بقضايا الحفظ في جهود ما وراء البيانات Working Group on Preservation Issues of Metadata التابعة لمجموعة المكتبات البحثية RLG تقريراً ختامياً يتضمن توصية بستة عشر عنصراً من عناصر البيانات التي تدعم حفظ الأصول الرقمية (http://www.rlg.org/preserv/presmeta.html)، ولما تبين أن معظم هذه العناصر من قبيل عناصر ما وراء البيانات الفنية المتصلة بالصور الثابتة ، أدمجت بعد ذلك في مواصفة المنظمة الوطنية لمعايير المعلومات والجمعية الدولية لإدارة المعلومات والصور التي سبقت الإشارة إليها.

وقد ظهر تطوران آخران خلال هذه الفترة وكان لهما تأثيرهما القوي في جدوى تحليل ما وراء البيانات المتصلة بعملية الحفظ ، وقد تمثل أول هذين التطورين في المسودة الأولى التي صدرت في عام ١٩٩٩م من النموذج المرجعي لنظام المعلومات الأرشيفية المفتوح Open Archival Information System(OAIS)[١]، والذي تبنت تطويره اللجنة الاستشارية لنظم البيانات الفضائية التابعة لوكالة ناسا للفضاء NASA's Consultative Committee for Space Data Systems ، ومع أن هذا النموذج يقدم إطاراً فكرياً للنظم

(١) Consultative Committee for Space Data Systems, "Reference Model for an Open Archival Information System (OAIS)," CCSDS ٦٥٠,٠-R-٢, Red Book, July ٢٠٠١, available at http://www.ccsds.org/documents/pdf/CCSDS-٦٥٠,٠-R-٢.pdf Accessed ٥ July ٢٠٠٢

الأرشيفية المعنية بالحفظ طويل الأجل للبيانات الرقمية ، إلا أنه لم يعتمد كمعيار دولي حتى الآن ، ولا تزال المنظمة الدولية للتقييس ISO تدرسه تمهيدًا لإقراره ليصبح أحد معاييرها المعتمدة.

أما التطور الثاني فيتمثل في إطلاق عديد من المشروعات الضخمة لأرشفة مصادر المعلومات الرقمية التي تبنت النموذج المرجعي لنظام المعلومات الأرشيفية المفتوح - المشار إليه في الفقرة السابقة (أو على الأقل درست إمكانية تبني هذا النموذج) . من هذه المشروعات : مشروع مكتبة الإيداع الأوربية الشبكية Networked European Deposit Library - المعروف اختصاراً " ندليب NEDLIB"، والذي يسعى إلى توسعة نظم الإيداع الوطنية لتشمل الأعمال الرقمية ، كذلك هنالك مشروع " سيدارز " CEDARS (١)في المملكة المتحدة، ومشروع المكتبة الوطنية لأرشيف باندورا الأسترالي National Library of Australia's PANADORA archive .

ولأن مجتمع المكتبات حاول أن يطبق خبرته العملية المبكرة في جهود الأرشفة كما يجسدها إطار نظام المعلومات الأرشيفية المفتوح ، فقد نشرت مجموعات عناصر ما وراء البيانات التي استخدمتها تلك المشروعات وغيرها، واستخدمت كأساس لمجموعة من الدراسات التحليلية اللاحقة . كذلك أصدرت مجموعة العمل الخاصة بما وراء بيانات الحفظ التابعة لكل من (أو سي إل سي) ومجموعة المكتبات البحثية، مجموعة توصيات خاصة بالجوانب الدلالية لما وراء بيانات الحفظ اعتماداً على عناصر البيانات

(١) مشروع دراسة واستكشاف قضايا الحفظ الرقمي [المترجم].

المستخدمة في المشروعات المشار إليها في الفقرة السابقة (نيدليب ، وسيدارز ، وباندورا) وقامت بتنظيمها ضمن إطار نظام المعلومات الأرشيفية المفتوح . ومع أن العناصر المقترحة لم تختبر جدواها بعد ، إلا أنها ستحظى من غير شك بدعم مواصفات أخرى أكثر نضجاً ، كما أنها أظهرت أن هناك بالفعل توجهاً محتملاً للعمل المستقبلي في مجال ما وراء بيانات الحفظ .

وفي سياق نموذج نظام المعلومات الأرشيفية المفتوح هنالك حزمة المعلومات information package التي هي عبارة عن مجموعة محتويات bundle of content ، مع ما وراء بيانات تستخدم لتقديم المحتوى القابل للأرشفة إلى مستودع البيانات واسترجاعه منه ، فضلاً عن إدارة هذا المحتوى الذي يضمه المستودع . وعادة ما تأتي حزم المعلومات في تجميعات مؤلفة من أربعة أنواع من المعلومات على النحو التالي : المعلومات الخاصة بالمحتوى content information ، والمعلومات الخاصة بوصف عملية الحفظ preservation descriptive information ، والمعلومات الخاصة بعملية تعبئة أو تجميع المعلومات packaging information ، والمعلومات الوصفية descriptive information . وتعني التوصية التي تبنتها مجموعة العمل الخاصة بما وراء بيانات الحفظ التابعة لـ (أو سي إل سي)،ومجموعة المكتبات البحثية بالفئتين الأوليين فقط من هذه الفئات الأربعة، وبحيث توصف عناصر البيانات في كلتيهما من خلال أربع خصائص هي : الاسم ، والأصل (بمعنى الإشارة إلى المشروع الذي تداول هذا العنصر، وهل يتم تداوله في مشروع سيدرز أم مشروع نيدليب،

أم مشروع المكتبة الوطنية الأسترالية، أو أن هذا العنصر أقرته مجموعة العمل نفسها)، والتعريف، والغرض منه . ثم يتبع ذلك ببعض الأمثلة على نوعية البيانات التي تقدم . ويلاحظ في هذا الصدد أن بعض العناصر تفتت إلى تركيبات من العناصر الفرعية التابعة لها حيث تعطى القيم الفعلية للبيانات عند آخر مستوى يقع فيه العنصر الفرعي nested subelement.

ويلاحظ كذلك أن المعلومات الخاصة بالمحتوى في سياق نموذج نظام المعلومات الأرشيفية المفتوح تتألف من كل من كيان بيانات المحتوى نفسه the content data object أو كيان المحتوى ، الذي عبارة عن مجموعة البتات bit Stream الفعلية التي يتم حفظها ، والمعلومات الخاصة بعملية التمثيل أو ما وراء البيانات الخاصة بذلك ،وهذه الفئة الأخيرة تتكون من العناصر التي تصف كيان المحتوى والعناصر التي تصف بيئتي المكونات المادية والبرمجية اللازمة لعرض الكيان أو إتاحة الوصول إليه . وتوصي مجموعة العمل _ المشار إليها سابقاً _ في تقريرها بثلاثة عشر عنصراً لوصف كيان المحتوى ، نتناولها باختصار في الفقرة التالية .

العناصر الوصفية لكيان المحتوى :

يعرف عنصر وصف الملف بأنه " أوصاف فنية للملف أو الملفات التي يتألف منها كيان بيانات المحتوى " ويعتبر هذا العنصر بمثابة الفئة الدلالية لما وراء البيانات الفنية . وهناك تعليق يشير إلى أن بعض التطبيقات قد تحتاج إلى تقسيم هذا العنصر إلى بيانات صغيرة جداً granular data ، ويشير كذلك إلى إمكانية استخدام عناصر [مواصفة] " ما وراء البيانات الفنية

للصور الرقمية الثابتة " مع صيغ كيانات محددة ، ومن العناصر الأخرى التي يمكن وضعها تحت مظلة ما وراء البيانات الفنية عنصرا النوع البنائي structural type والحجم .

وتوصف السمات الأساسية لكيان بيانات المحتوى عبر عنصرين : أولهما عنصر " الوظيفية Functionality " الذي يستخدم للإشارة إلى الخصائص الوظيفية أو الشكلية والحسية look and feel للكيان ، وثانيهما عنصر " وصف المحتوى المعروض Description of Rendered Content " الذي يصف الكيفية التي ينبغي أن يظهر بها الكيان للمستفيد ، وتحسباً للحاجة إلى تحويل ملفات البيانات من صيغها الأصلية المتقادمة إلى صيغ مادية جديدة ، فقد خصصت ثلاثة عناصر تتعامل مع هذا الإجراء ، أولها يعرف " بميسرات الإتاحة أو الوصول Access Facilitators " ويهدف من ورائه إلى بيان " المعينات والميسرات aids and facilitators " التي ينبغي أن تؤخذ بعين الاعتبار عند عملية الحفظ ، مثل مؤشر الوقت time index المرتبط بالمقطع المتحرك . وثانيهما هو عنصر "السمات الجوهرية Significant Properties " الذي تسجل فيه سمات كيان المحتوى التي ينبغي أن تحفظ خلال إجراءات الحفظ. أما العنصر الثالث فهو عنصر " المراوغات Quirks " - ذلك المصطلح المقتبس من المكتبة الوطنية الأسترالية ، والذي يوثق من خلاله ما قد ينتج عن إجراءات الحفظ من فقدان لأي من جوانب الوظيفية أو من تغيير في الخصائص " الشكلية والحسية " .

وهناك عنصر " وصف النموذج المجرد الأساسي Underlying Abstract Form Description " الذي يفترض أن يشتمل على شرح - مكتوب بلغة

يفهمها البشر - للكيفية التي تفسر بها مجموعات البتات المحفوظة حفظاً تاريخياً أو المؤرشفة حتى يمكن عرض الكيان، وهناك على نحو مشابه عنصر " البنية التحتية الفنية للكيان المركب Technical Infrastructure of Complex Object " الذي يستخدم لوصف البنية الداخلية للكيانات متعددة الأجزاء ، مثل صفحات الوب التي تشتمل على روابط وملفات ضمنية . فضلاً عن ذلك هنالك عناصر أخرى مثل : عنصر " موانع الإتاحة Access Inhibitors " الخاصة بتشفير البيانات ووضع العلامات المائية وغيرها من آليات الحماية الأخرى ، وعنصر " متطلبات التجهيز " ، وأخيراً عنصر " التوثيق " .

أما العناصر المتبقية في فئة " معلومات التمثيل " فهي مجموعة من عناصر "الوصف البيئي Environment Description " وتقسم في ثلاث فئات ، تضم أولاها العناصر التي تصف البرامج المطلوبة لعرض الكيان أو إتاحته ، وتضم الثانية العناصر التي تصف نظم التشغيل التي تعمل تحتها تلك البرامج ، وتضم الثالثة العناصر التي تصف العتاد البرمجي أو المادي المطلوب . ونظراً لأن معظم الكيانات يمكن أن تعرض أو تتاح عبر بيئات متعددة تتضمن توليفات متعددة داخل تلك الفئات الثلاثة ، ونظراً كذلك لتتابع ظهور كل من الإصدارات الجديدة، والقديمة المتوافقة من المكونات المادية والبرمجية، يصعب التحكم في عدد قيم عناصر البيانات هذه والحفاظ على استمرارها، من هنا يتحتم على جهود التنفيذ الفردي أن تقرر متى يجب تسجيل البيانات في هذه الفئات بشكل شامل، أو متى يكتفى بتسجيل الحد الأدنى من المواصفات، وما هي المواصفات الحالية المفضلة، أو بعض

المجموعات الفرعية الأخرى لقيم البيانات .

معلومات وصف الحفظ :

تعد فئة المعلومات المتصلة بوصف عملية الحفظ هي الفئة الثانية الكبرى من ما وراء البيانات التي غطتها خطة مجموعة العمل ، المشار إليها من قبل. وقد جاءت العناصر في هذه الفئة في أربع فئات فرعية على النحو التالي : المعلومات المرجعية Reference information ، ومعلومات السياق Context information ، ومعلومات المنشأ أو المصدر Provenance information ، ومعلومات مدى الثبات Fixity information . حيث تستخدم المعلومات المرجعية لأغراض الوصف أو التعريف ، ومن ثم تحتوي على معرفات عالمية global identifier ، ومعرفات محلية خاصة بالأرشيف ، بالإضافة إلى بعض أساليب الإحالة إلى تسجيلة ما وراء البيانات الوصفية الموجودة للمصدر . أما معلومات السياق فتوثق الهدف الكامن من وراء إنشاء كيان المحتوى ، ومدى العلاقة بين هذا الكيان والكيانات الأخرى ، وتتوافر في هذا الصدد عناصر خاصة بتوثيق الكيفية التي يربط بها هذا الكيان بوسائط العرض manifestations الأخرى المرتبطة به (مثال ذلك الإصدارة السابقة للعمل نفسه)، وتوثيق الكيفية التي يرتبط بها هذا الكيان فكرياً مع الكيانات الأخرى (مثال ذلك، ارتباطه بالمواد الأخرى في المجموعة نفسها) .

أما فئة معلومات المنشأ فتوثق تاريخ معلومات المحتوى ، وقد قسمت مجموعة العمل في تقريرها هذه الفئة من المعلومات إلى خمس فئات [فرعية]، تتعامل أربع منها مع الجوانب المتصلة بدورة حياة الكيان والمراحل التي مر بها وقد جاءت في أربع مراحل : مرحلة الميلاد Origin، ثم مرحلة النشأة

أو ما قبل البلوغ Pre-Ingest ، ثم مرحلة التعبئة أو البلوغ Ingest، وأخيراً مرحلة الحفظ الأرشيفي Archival Retention ، بحيث توصف في كل مرحلة من هذه المراحل الأحداث التي أحاطت بالكيان من حيث : ماذا حدث للكيان بالضبط ، ومتى حدث هذا ، ومن كان مسؤولاً عن الحدث ؛ إذ يمكن على سبيل المثال أن يقوم الأرشيف بإعادة وضع المادة في صيغة أخرى في الوقت الذي تكون فيه في طريقها إلى الحفظ التاريخي أو الأرشفة ، وفي هذه الحالة يسجل هذا الحدث على أنه حادثة مرتبطة بمرحلة التعبئة an Ingest event . وإذا كانت هذه الفئات الأربع تتصل بدورة حياة الكيان ، فتأتي الفئة الخامسة المسماة " إدارة الحقوق Rights Management " لتحديد " الاستخدامات القانونية لكيان المحتوى".

أما معلومات " مدى الثبات " فتقدم من خلالها المعلومات التي تساعد على التثبت من عدم تغير الكيان بطرق غير قانونية أو غير موثقة . وتحدد التوصية عناصر تتصل بالجوانب التالية : نوع المصادقة Authentication Type، وإجراءاتها ، وتاريخها ، ونتيجتها . فإذا كان أسلوب المصادقة يعتمد مثلاً على نظام مجموع التدقيق اللوغاريتمي المسمى " MD٥ checksum "(١)، فإن نوع المصادقة في هذه الحالة يكون " MD٥ " ، أما عنصر " إجراءات المصادقة " فيشار فيه أو يتضمن توثيقاً يشير إلى وصف الإجراء الذي تم عن طريقه تكوين سلاسل مجموع التدقيق والمسماة " MD٥ strings

(١) أحد الأساليب اللوغاريتمية التي يعتمد عليها في تدقيق ملفات البيانات لاختبار صحة البيانات واكتمالها ، ويرجع اختراع هذا الأسلوب إلى رونالد إل ريفست Ronald L. Revist من معهد ماسوشتس للتقنية [المترجم] .

" ، وأما نتيجة المصادقة فيمكن أن تكون هي السلسلة التي يرد فيها مجموع التدقيق نفسه .

وتجدر الإشارة إلى أن عناصر " مدى الثبات " لا تبين سوى طريقة واحدة تدعم من خلالها توصيات مجموعة العمل لمركز أو سي إل سي ومجموعة المكتبات البحثية النموذج المرجعي لنظام المعلومات الأرشيفية المفتوح . وفي سياق هذا النظام ، لم يفترض أن المجتمعات المستقبلية سوف تستوعب التقنيات الجارية أو ستتاح لها إمكانية الوصول إلى هذه التقنيات . فالإشارة ببساطة إلى نوع المصادقة (كما هو الحال في نظام MD٥) لا يعد كافياً بدون توثيق للكيفية التي يتم بها عمل هذا النظام من نظم المصادقة ، وبالمثل لا يكفي عند تقديم المعلومات المرجعية تسجيل نوع المعرف (مثل الترقيم الدولي الموحد للكتب ISBN) دون شرح للكيفية التي يتم بها تكوين هذا الترقيم . وعليه يمكن القول إن الحاجة إلى تسجيل التوثيق الداعم [للإجراءات التي تتخذ] يعد من الأمور الأساسية بالنسبة لنظام المعلومات الأرشيفية المفتوح ، وهو ما يتواجد دائماً في سياق كل التوصيات التي تطرحها مجموعة العمل .

توصيات وقضايا :

ما أشبه مجموعات عناصر ما وراء البيانات التي أوصت بها مجموعة العمل لـِ (أو سي إل سي) ومجموعة المكتبات البحثية بنظام المعلومات الأرشيفية المفتوح نفسه من حيث الغرض منها ؛ حيث قصد من هذه العناصر أن تكون بمثابة إطار أو نظام يعتمد عليه عند تطوير تطبيقات ما

وراء بيانات معينة ، وليس كمواصفة رسمية لإحدى خطط ما وراء البيانات . ومن الملاحظ أن التطبيقات المختلفة تحتاج إلى تطبيق عناصر معينة على مستويات مختلفة من التفصيل الشديد؛ فقد يلبي عنصر وحيد من عناصر وصف الملف - على سبيل المثال - متطلبات تطبيق واحد ، بينما قد يتطلب تنفيذ تطبيق آخر مجموعة من عناصر ما وراء البيانات الفنية . وأنه بقدر ما تكون [خاصية] أنواع البيانات للعناصر غير محددة، يكون للتطبيق حرية تنفيذ عملية التوثيق ، حيث يمكن أن يسجل كجملة نصية أو كمؤشر pointer ، وبالمثل فعندما لا تتضمن المواصفة الإشارة إلى أية لغات مقيدة أو لا توصي بلغات محددة ، لا تكون هنالك أي تركيبة نقل محددة .

ومع محاولة التطبيقات الفعلية للتعامل مع التوصيات كإطار عمل لما وراء بيانات الحفظ الخاصة بها، أثيرت قضايا معينة ، منها القناعة بأن عددًا من الكيانات في الأرشيف الفعلي للحفظ تتطلب أن تكون مقيدة أو منضبطة على مستويين : أولهما ككيانات منطقية logical entities ، وثانيهما كمجموعات من الملفات الفردية المكونة لهذه الكيانات المنطقية؛ فصفحة الوب - على سبيل المثال- تعد كياناً منطقياً مكوناً من العديد من الملفات المادية (المتمثلة في التتابع المسلسل للبتات) ، مثل ملف في صيغة لغة ترميز النص الفائق ، وملف آخر أو أكثر من ملفات الصور في صيغة جيف ، كما أن الكتاب المحول إلى صيغة رقمية كمجموعة من صور لصفحات في صيغة تيف Tiff يعد مثالاً آخر على ذلك؛ حيث نجد في بعض الأحيان ولأهداف معينة أن المادة التي يجب أن تخضع للوصف والإدارة هي الكتاب نفسه،

وبما يتصل به من خصائص ، كالمؤلف والعنوان والتركيبة المعقدة. ولكن في مواقف أخرى ولأهداف معينة كذلك ـ كما هو الحال عند تجديد الوسائط التي يحمل عليها الكتاب وتحويل البيانات أو نقلها الى الصيغ الأحدث ـ ينبغي أن تخضع كل صورة من الصور الممثلة للصفحات والتي كانت في صيغة تيف للإدارة على نحو مستقل . ومع أن مجموعة العمل لِـ (أوسي إل سي) ومجموعة المكتبات البحثية أقرت بهاتين الطبقتين ، إلا أنها لم تقدم إطاراً صريحاً يتم بناءً عليه تصنيف عناصر الوصف التي تخص كلاً منها، وكيف يمكن ربط كل منها بشكل معماري أو فني.

وهناك قضية أخرى وهي أن عديدًا من عناصر ما وراء البيانات المقترحة يبدو أنها تحتاج إلى تحليل تفصيلي للمحتوى المحفوظ حفظاً تاريخياً [الذي تم أرشفته]، وقد يكون من غير الممكن إجراء ذلك عملياً ، أو تنفيذه في جميع المواقف، وخاصة في سياق مشروعات الحفظ الكبيرة . وعلى أية حال ، يمكن القول إن الاهتمام المكثف والاستثمار المتزايد في مشروعات الحفظ طويل الأجل للمجموعات الرقمية سوف يكفل لخطط ما وراء بيانات الحفظ استمرارية تطويرها.

READINGS :

- Digital Audio-Visual Preservation Prototyping Project of the Library of Congress (home page). Available at http://www.loc.gov/rr/mopidavprot/avprhome.html. Accessed ٥ July ٢٠٠٢.

This page links to the Project Document Menu, which contains links to background documents, planning documents, and documents about metadata (http://lcweb.loc.gov/rr/mopic/avprot/avlcdocs.html).

These include Extension Schemas for the Metadata Encoding and Transmission Standard (see chapter ١٦), which in turn link to schemas and data dictionaries defining technical metadata for audio and video.

- Library of Congress Digital Library Development, Core Metadata Elements, available at http://www.loc.gov/standards/metadata.html.

  This table of administrative, descriptive, and structural metadata elements used by the Library of Congress is interesting because it categorizes elements according to several vectors, including their function (access management, presentation, preservation, etc.) and the level of object to which they apply (set, aggregate, primary object, etc.).

- OCLC/RLG Preservation Metadata Working Group (home page). Available at http://www.oclc.org/research/pmwg/. Accessed ٥ July ٢٠٠٢.

  Check the "Documents" link here for the Working Group report discussed earlier and any successor documents.

الفصل السابع عشر

ما وراء البيانات البنائية

تُعنى ما وراء البيانات البنائية بوصف التنظيم الداخلي لمصادر المعلومات؛ ذلك أن المصادر المنطقية في البيئة الرقمية عادة ما تتكون من عدة ملفات مادية، ومن ثم تقوم هذه الفئة من ما وراء البيانات بربط هذه الملفات المادية بعضها مع بعض من ناحية، وربطها ببنية الكيانات المنطقية من ناحية أخرى ؛ فإذا كان لدينا - على سبيل المثال - كتاب في مئة صفحة، ثم حول إلى صيغة رقمية، ولتكن صيغة تيف TIFF، فستكون لدينا مئة صورة كملفات تيف، إذ إن كل صورة تمثل صفحة واحدة من صفحات هذا الكتاب . وهنا يبرز دور ما وراء البيانات البنائية في تحديدها لأي من ملفات تيف على أنه هو الذي يمثل الصفحة الأولى من الكتاب ، وأيها يمثل الصفحة الثانية، وهكذا، وفي تحديدها كذلك أن الصفحات من ١ - ١٤ تمثل الفصل الأول ، وأن الصفحات من ١٥- ٢٤ تمثل الفصل الثاني - مثلاً - وهكذا، وبناء على هذه المعلومات ، يمكن كتابة أحد التطبيقات المباشرة لتنفيذ وظائف محددة، كعرض قائمة محتويات الكتاب، وإتاحة الفرصة للقارئ للانتقال مباشرة إلى إحدى صفحات الكتاب، كالصفحة العاشرة، أو الفصل الثاني مثلاً، وكذلك إتاحة إمكانية تصفح الكتاب من بدايته إلى نهايته أو العكس .

كذلك يمكن لما وراء البيانات البنائية أن تتيح إمكانية الربط بين أشكال التمثيل المختلفة للمحتوى الفكري نفسه ، حيث يمكن أن يكون بالنسبة لصفحات الكتاب - الذي أشرنا إليه كمثال في الفقرة السابقة - والتي تمثلها ملفات تيف صور أخرى ، ولكن في صيغة جي بيج JPEG ، أو صور أخرى مصغرة thumbnail، وهنا تبرز أهمية ما وراء البيانات مرة أخرى في ربطها

هذه الصيغ الثلاثة للملفات بعضها مع بعض، ومن ثم يمكن للمستفيد أن يختار أسلوب العرض الذي يناسبه.

وهنالك استخدام ثالث لما وراء البيانات البنائية في موقف آخر يتمثل في إمكانية ربط مقاطع معينة في أحد التسجيلات الصوتية بالمقاطع المناظرة لها في الملف المطبوع أو الملف الموسيقي [للعمل نفسه] .

وبقدر ما تكتسب ما وراء البيانات البنائية أهميتها في عرض [مصادر المعلومات الرقمية] ، تتجسد أهميتها كذلك في إدارة هذه المصادر وحفظها ؛ فالمستودع الذي يتولى مهام اختزان المصادر الرقمية ، يحتاج إلى معرفة الملفات التي تتكون منها الكيانات حتى يتمكن من تنفيذ أية إجراءات تتصل بهذه الكيانات المنطقية ، مثل عمليات الإضافة ingestion (١)، وكتابة التقارير ، والبث أو التوزيع.

وتعد نظم ترميز النصوص في سياقات أخرى بمثابة ما وراء بيانات بنائية ؛ فعندما يرمز محتوى نصي - على سبيل المثال - بصيغة لغة الترميز المعيارية العامة SGML ، أو بصيغة لغة الترميز الموسعة XML اعتماداً

(١) ingestion من المصطلحات المتداولة حديثاً في مجال ما وراء البيانات ليعني التغذية أو التعبئة أو الإضافة أو حتى التزويد، أي نقل وإضافة بعض الملفات من مكان ما إلى ملفات الكيان الرقمي أو إضافة ما وراء بيانات من مكان ما إلى مكان آخر ، ولعل أول طرح لهذه الفكرة جاء في سياق أحد الأبحاث العلمية حول تطوير خادم عملاق لمعلومات البيانات العلمية(Scientific Data Information Super Server (SDISS ، وقد جاء النظام المقترح في سياق دراسة بعنوان : Automatic Metadata Ingestion for Supporting A Web-Based And Information Super Server Scientific Data للباحث Ruixin Yang ، حيث تناول فيها إجراءات التغذية والنظام المقترح للقيام بهذه المهمة [المترجم].

علـى إرشـادات مـبادرة ترميـز الـنص TEI Guidelines ، فمـن المـرجح أن يـشتمل ذلـك الترميـز علـى معلـومات بنائـية كتلك التـي تحـدد الفواصـل بـين الـصفحات، وعـناوين الفـصول، وغيـر ذلـك مـن الأجـزاء الرئيـسية للـنص . وتحقيقاً لهذه الخدمات كـان علـى خطـط ما وراء البيانات البنائية أن تشتمل علـى مواصفات مثل مواصفة إرشادات ترميز النص ، وتركيبة نشر الكتاب الإلكتروني المفتوحة the Open eBook Publication Structure التي طورها منـتدى الكـتاب الإلكترونـي المفـتوحOeBF، ومواصـفة دوك بـوك DocBook التـي طـورها نظـام المعلـومات الأرشـيفية المفـتوح OASIS . وسـوف نركـز حديثنا فـي هذا الفصل علـى الخطـط التـي لا تـشتمل علـى [قـواعد خاصـة] بالمحتوى ، أو التي تشتمل على مثل تلك القواعد ولكن بشكل اختياري.

صيغة إيفيكت **EFFECT** (تبادل المكونات والنصوص الإلكترونية) :

فـي بدايـة التـسعينات مـن القـرن العـشرين المـيلادي طـور " إلـسفير" ـ الناشر المعروف، صيغة تبادل المكونات والنصوص الإلكترونية Exchange Format for electronic components and Texts المعروفة اختـصاراً بـصيغة إيفـيكت EFFECT، بهـدف دعـم إجـراءات توزيـع الملفـات الإلكتـرونية مـن الناشـرين إلـى المكتبات وغيرها من نظـم الاستـضافة البعيدة . وفـي إطـار العمـل الخـاص بهـذه الـصيغة يـتم تجمـيع محـتوى الدوريـة الإلكتـرونية فـي مجمـوعات تعـرف بمجمـوعات البيانات datasets ، والتي قد تكون مبنية إما علـى أسـاس أعـداد الدوريـة issue-based ، أو علـى أسـاس مقالاتـها article-based . مـع ملاحظـة أن كل مجمـوعة مـن هذه المجمـوعات تكون في ملف مـستقل ينتهـي بالامـتداد " toc. " ، ويعـرف بالتالـي هكـذا " "dataset.toc"" ،

وداخل هذا الملف تقسم المحتويات طبقياً ، ويشار إلى كل مستوى بالتاج " tn_" بحيث يأخذ أعلى المستويات التاج " t٠_" ، ثم يأخذ المستوى الذي يليه التاج "t١_" ، ثم التاج " t٢_ " للمستوى التالي ، وهكذا . وقد خصص التاج " t٠_" وتوابعه لتعريف مجموعة البيانات نفسها ، بما في ذلك رقم هويتها في هذه المجموعة ، والإصدارة المستخدمة من صيغة إيفيكت ، وتاريخ إنشاء المجموعة ووقت إنشائها . أما التاج "t١_ " وتوابعه فقد خصصت لتعريف عنوان الدورية ، كما خصصت التاج " t٢_" لتعريف العدد ، والتاج " t٣_ " لتعريف المقالة المنشورة بالعدد وغير ذلك من المساهمات . وذلكم مثال لجزء من ملف بإحدى مجموعات البيانات dataset.toc ، استخدم للتعريف بإحدى المقالات المنشورة في مجلة " Brain research " ، في العدد الأول من المجلد رقم ٩٤٥ .

_tl	AAA٠٠٠٠١ ٠٠٠٦٨٨٨٦
_jn	Brain Research
_pu	Elsevier Science
_t٢	AAA٠٠٠٠١ ٠٠٠٦٨٩٩٦ v٠٩٤٥i٠١
_vl	٩٤٥
_is	١
_dt	٢٠٠٢٠٧٢٦
_t٣	AAA٠٠٠٠١ ٠٠٠٦٨٩٦ v٠٩٤٥i٠١ ٠٠١٢٣٤٥٦
_ii	[SIC]…
_ti	Increases in amino-cupric-silver staining of the supraoptic nucleus after sleep deprivation
_pg	١-٨

_mf [TIFF ٦,٠] ١.tif ٢.tif ٣.tif ٤.tif ٥.tif ٦.tif ٧.tif ٨.tif
_mf [raw ASCII] ١.raw ٢.raw ٣.raw ٤.raw ٥.raw ٦.raw ٧.raw ٨.raw

ومن الملاحظ على هذا المثال أن هناك ثلاثة تيجان تندرج تحت المستوى الثاني الذي يأخذ التاج " t٢_" الخاص بتعريف العدد ، وهذه التيجان هي : "v١_" وخصص للإشارة إلى رقم المجلد ، والتاج "is_" وخصص للإشارة إلى رقم العدد، والتاج "dt_" وخصص للتسمية الزمنية للعدد أو تاريخ نشره . وهناك كذلك التاجان "ti_" ، و "pg_" في المستوى الثالث الذي يأخذ التاج "t٣_" الخاص بتعريف المقالة ؛ حيث خصص أولهما - "ti_" - لتسجيل عنوان المقالة ، وخصص الآخر - "pg_" - لتسجيل الصفحات التي تشغلها المقالة ،وعلى هذا المنوال يمكن تعريف المؤلفين والموضوعات والمستخلص والكلمات الدالة وغيرها من العناصر التي تدخل في إطار ما وراء البيانات الوصفية .

أما التاج " mf_" (الذي يشير إلى وسيط العرض manifestation) فيخصص لتعريف الملفات التي تتألف منها المقالة ، حيث يلاحظ أن هناك إصدارتين أو صيغتين لتمثيل الملفات استخدمتا لتمثيل صفحات المقالة المشار إليها ، وهي الصفحات من ١ـ٨ ، حيث استخدمت صيغة تيف لتمثيل صور الصفحات ، وصيغة آسكي الخام "raw" ASCII لتمثيل الملفات النصية بدون ترميز . مع ملاحظة أن الأمر قد لا يقتصر على هاتين الصيغتين ، بل يمكن تمثيل الصيغ الأخرى مثل صيغة لغة ترميز النص الفائق HTML ، وصيغة لغة الترميز المعيارية الموحدة SGML ، وصيغة

بي دي إف PDF .

ومع أن صيغة إيفيكت هذه كانت قد صممت لدعم توزيع مطبوعات الناشر إلسفير في المقام الأول ، فقد تبناها ناشرون آخرون ، مع إجراء بعض التعديلات الطفيفة عليها . كما يشار كذلك في هذا الصدد إلى أنه عندما كانت طريقة "التحميل المحلي local loading " هي الطريقة الشائعة لتداول الدوريات الإلكترونية، كان للعديد من المكتبات سبل معالجة مهيأة لاستقبال وعرض محتوى الدوريات الإلكترونية المقدمة في ملفات مجموعات البيانات المعتمدة على صيغة إيفيكت .

مشروع إي بايند أو التجليد الإلكتروني **EBIND** :

مشروع بيركلي للتجليد الإلكتروني the Berkely Electronic Binding Project ، أو ما يعرف اختصاراً " إي بايند EBIND " ، محاولة مبكرة (١٩٩٦م) لتقنين ما وراء البيانات البنائية الخاصة بالمصادر المحولة إلى الصيغ الرقمية في شكل صور للصفحات ، ويقوم هذا المشروع على أساس تعريف ما وراء البيانات البنائية باستخدام أحد معرفات نوع الوثيقة في صيغة اللغة المعيارية الموحدة للترميز SGML DTD والمعتمد بشكل غير مقيد على معرف نوع الوثيقة الخاص بمبادرة ترميز النص TEI DTD . ومن ثم فإن معرف نوع الوثيقة الخاص بهذا المشروع " وهو Ebind DTD " يأتي مشابهاً لذلك المعرف الخاص بمبادرة ترميز النص ، حيث يشتمل على قسم الترويسة (<ebindheader>) لما وراء البيانات الببليوجرافية الوصفية للمصدر ، يليه قسم محتوى المصدر ، وهو مُقسَّم بدوره إلى ثلاثة أقسام

فرعية : قسم خاص بالعناصر التمهيدية[١] (<front>) ، وقسم خاص بالمتن (<body>) ، ثم قسم خاص بالعناصر الختامية back matter (<back>) .

وكما هو الحال كذلك في مبادرة ترميز النص ، يتم في سياق صيغة التجليد الإلكتروني الإشارة إلى الأقسام البنائية من خلال العناصر المسماة عبر التاج <div> - أي قسم ، وفي حال تسمية هذه الأقسام البنائية بأسمائها، مثل عناوين الأجزاء أو الفصول ، فإن هذه المعلومات تسجل من خلال العناصر المسماة عبر التاج <head> ، أي عناوين رئيسية، بحيث يمكن أن تمثل الأقسام التمهيدية الخاصة بأحد الكتب على النحو التالي :

```
<front>
<div٠ type="titlepage">
<div٠ type="preface">
<head>Preface to the second edition</head>
<div٠ type= "'contents ">
<head>Table of Contents</head>
</front>
```

ومن الملاحظ أن هناك ثلاثة أقسام في هذا المثال تشكل الأجزاء التمهيدية للكتاب ، وهي : صفحة عنوان بدون اسم محدد، ومقدمة لها عنوانها، وقائمة محتويات لها عنوانها كذلك[٢]. وعندما تتواجد صور

---

(١) تعرف في سياق مجال الفهرسة بالقوادم ، ويعبر عنها بمصطلحات أخرى منها preliminaries or preliminary matter [المترجم].

(١) يشار في المثال إلى وجود صفحة عنوان ، ، إلى مقدمة بعنوان " Preface to the second edition ، وإلى قائمة محتويات "عنوانها " Table of Contents " [المترجم].

الـصفحات الخاصــة بهذا العمـل، فـإن تمثيل هـذه الـصفحات يـشار إلـيه فـي العنـصر الفرعـي <image> - أي صـورة، وذلـك داخـل العنـصر الرئيـسي <page> - أي الـصفحة، كمــا تـشمل خـصائص العنـصر الفرعـي المـسمى <image> - أي صـورة الخـصائص الـثلاثة التاليـة[١]: خاصـية "entityref" مرجع الكيان، وخاصية "idref" مرجع التعريف لتعريف ملفات الـصور، وخاصـية رقـم التـتابع "seqno" التـي يـسجل بهـا رقـم التـتابع الفـريد لملـف الصورة داخل المصدر، وخاصية الرقم الأصلي "nativeno" التي يسجل بها رقم الـصفحة الـذي يظهـر علـى الـصفحة المـرقمنة أو المحـولة إلـى صـيغة رقمية، وذلك على النحو التالي :

```
<front>
<div٠ type="titlepage">
<page><image entityref="QA٠٠٠٠١" seqno="١"></page>
<div٠ type="preface">
<head>Preface to the second edition</head>
<head><image entityref="QA٠٠٠٠٢" seqno="٢"
nativeno="i"></page>
<div٠ type= "'contents ">
<head>Table of Contents</head>
<page><image entityref="QA٠٠٠٠٣" seqno="٣"
nativeno="ii"></page>
```

(٢) ذكر مسميات الخصائص هو من اجتهاد المترجم ، وما ذكر في النص الأصلي هو مجرد الاسم المختصر أو التاج ، مثل seqno ، و idref [المترجم].

```
<page><image entityref="QA٠٠٠٠٤" seqno="٤"
nativeno="iii"></page>
</front>
```

إن بساطة صيغة إي بايندر هي أبرز ما يميزها، حيث يمكن بسهولة إنشاء ملف إي بايندر في صيغة لغة الترميز المعيارية الموحدة Ebind SGML ، وذلك عن طريق إحدى البرمجيات بعد تجهيز الحد الأدنى من البيانات المدخلة في نموذج مخصص لهذا الغرض ، أو ما يعرف بنماذج إدخال البيانات Worksheet . يضاف إلى ذلك أن تركيبة "إي بايندر" تتيح قدراً كبيراً من المرونة نتيجة استخدامها لعناصر الأقسام <div> التي يتم تضمينها عناصر فرعية من ناحية، وبسبب قدرتها على استيعاب أي مستوى من مستويات التقسيمات الهرمية أو الطبقية من ناحية أخرى ، هذا بالإضافة إلى أنه في حال توافر نسختين للنص الكامل، إحداهما مطبوعة والأخرى لصور الصفحات ، يمكن في هذه الحالة إدراج النص في ملف "إي بايند" مع ترميزه ترميزاً مناسباً في صيغة مبادرة ترميز النص TEI .

معيار نهضة أمريكا الثاني MOA٢ ، ومعيار ترميز ما وراء البيانات ونقلها : METS

في عام ١٩٩٥م بدأت جامعتا ميتشيجن وكورونيل مشروعاً أطلق عليه اسم "تقدم أو نهضة أمريكا making of America " وهو ما يعرف اختصاراً بـ "MOA"، لرقمنة المنفردات والمسلسلات الخاصة بالتاريخ الاجتماعي الأمريكي منذ فترة ما قبل الحرب إلى فترة

إعادة الإعمار . وفي عام ١٩٩٧م بدأ أعضاء ائتلاف المكتبات الرقمية مشروعاً آخر أسموه " نهضة أمريكا الثاني making of America II " أو اختصاراً MOA٢ ، بتمويل من الوقف الوطني للإنسانيات ، وقد خطط لهذا المشروع ليكون تجربة لبناء مجموعة متكاملة غير مركزية من المواد الأرشيفية الرقمية التي يقصد بها في سياق هذا المشروع المذكرات اليومية ودفاتر الحسابات ledgers وألبومات الصور، مع الاستبعاد الكامل للكتب والمنفردات . وقد اشترك في إنجاز هذا المشروع جامعات: بيركلي ، وكرونيل ، وبنسلفانيا الحكومية ، وستانفور ، بالإضافة إلى مكتبة نيويورك العامة ، على أن تقود جامعة كاليفورنيا في بيركلي هذا المشروع . وكان من أهم ثمرات هذا المشروع معيار ما وراء بيانات جديدة ، يعرف باسم هذا المشروع نفسه، وهو معيار " نهضة أمريكا الثاني MOA٢ ".

وبينما كان مشروع التجليد الإلكتروني أو "إي بايند" مصمماً في الأساس لدعم وظائف العرض كتقليب الصفحات ، اتخذ مشروع " نهضة أمريكا الثاني " اتجاهاً أكثر شمولية ؛ إذ توجه لدمج كل من ما وراء البيانات الوصفية والإدارية والبنائية في معرف واحد من معرفات نوع الوثيقة في صيغة لغة الترميز الموسعة XML DTD ، ومن ثم جاء هذا المعرف في أربعة أقسام : أولها قسم ما وراء البيانات الوصفية ، ويستخدم للإحالة إلى إحدى تسجيلات ما وراء البيانات الخارجية ، أو يشتمل على ما وراء بيانات وصفية ضمنية ، وثانيها قسم ما وراء البيانات الإدارية ، ويقدم في سياقه ثلاثة أنواع من البيانات هي : ما وراء البيانات الفنية التي تختص بتعريف

الملف من حيث إنشائه وخصائصه المادية ، ومعلومات عن حقوق الملكية الفكرية الخاص بالمصدر الموصوف ، ثم معلومات عن المصدر الأصلي للكيان الرقمي ، أما ثالث الأقسام فهو قسم ما وراء البيانات البنائية ، ويعني بتنظيم الملفات في أماكنها المنطقية باعتبارها من أجزاء الوثائق المركبة ، وأخيراً يأتي القسم الرابع المسمى " سجل الملف file inventory " ويعني بتجميع كل الملفات الخاصة بنسخة معينة من نسخ الكيان الأرشيفي (مثال ذلك النسخة في صيغة جي بيج أو النسخة المصغرة) . ومن الملاحظ أن القائمة الخاصة بكل ملف مستقل قد تتضمن مؤشرات أو موجهات pointers إلى معلومات ذات صلة في قسم ما وراء البيانات الإدارية أو أنها قد تشتمل على ما وراء بيانات إدارية متضمنة داخلها .

وفي بداية عام ٢٠٠١م، نظم اتحاد المكتبات الرقمية ورشة عمل لمناقشة مواصفة ما وراء البيانات الخاصة بمشروع " نهضة أمريكا الثاني " وكيفية تعديلها لتلبية المزيد من الاحتياجات ، وقد أقرت المجموعة المنظمة لورشة العمل نجاح خطة ما وراء البيانات في سياق هذه المرحلة التجريبية Testbed لهذا المشروع - أي نهضة أمريكا الثاني ، إلا أنها انتقدت افتقار هذه الخطة لإمكانات الربط الخارجي ، وأنها صممت أساساً للعمل مع مجموعة محددة من مصادر المعلومات ومن ثم لا تتناسب وطبيعة الوسائط الصوتية والمرئية وغيرها من الوسائط المعتمدة على عنصر الوقت time-dependent. وبناءً على ما انتهت إليه ورشة العمل من ملحوظات تتصل بهذه الخطة ، بدأ العمل في تطوير صيغة معدلة ، اتفق على تسميتها " بمعيار تكويد ما وراء البيانات وتناقلها Metadata Encoding and Transmission

Standard "، أو اختصاراً : ميتس METS ، وقد ظهرت نسختها التجريبية في صيغة لغة الترميز الموسعة في صيف ٢٠٠١م، ثم ظهرت إصدارتها الأولى - version١,٠ - في فبراير ٢٠٠٢م، وقد تولى مكتب تطوير شبكة مكتبة الكونجرس ومعايير مارك the library of Congress Network Development and MARC Standards Office ، رعاية هذه الخطة واستضافة موقعها الرسمي على الوب (http://www.loc.gov/standard/mets/)، وفي مايو من العام نفسه (٢٠٠٢م) أعلن عن لجنة تحرير المعيار على أن تعمل بمثابة هيئة تتولى إدارة مراجعته.

يتكون معيار ميتس METS - شأنه شأن معيار نهضة أمريكا الثاني - من الأقسام الرئيسية الخاصة بكل من : ما وراء البيانات الوصفية ، وما وراء البيانات الإدارية ، وما وراء البيانات البنائية وسجل الملف، غير أن هناك فارقاً بينهما وهو اشتمال هذا المعيار على قسمين أخريين، أحدهما، وهو القسم الرابع والذي يعرف بقسم الترويسة، يعني بتسجيل ما وراء البيانات الخاصة بملف ميتس نفسه، والآخر، وهو القسم الخامس، ويعنى بتسجيل التصرفات المرتبطة بالكيان، كما أن هناك فرقاً كبيراً آخر يتصل بتصميم هذا المعيار، وهو أن خطة ميتس لا تحدد أية عناصر لما وراء البيانات الوصفية أو الإدارية، وإنما تستخدم بدلاً من ذلك أحد أسلوبين، إما الإشارة أو الإحالة إلى تسجيلة خارجية لما وراء البيانات، وإما تضمين ما وراء البيانات التي تستقى من اسم مميز آخر namespace (ليس من خطة ميتس non-METS) داخل أحد العناصر المغلِفة wrapper element . ومن الملاحظ أن أسلوب استخدام خطة التوسعة أو خطة ما وراء البيانات المعرفة

خارجياً في صيغة لغة الترميز الموسعة يلقي تشجيعاً بالنسبة لكل من ما وراء البيانات الوصفية والإدارية ، وبهذه الطريقة يستطيع ميتس أن يدعم عمل مجموعات المعايير الأخرى ، ويتجنب إشكالات الصيانة التي قد تنشأ إذا حددت عناصر ما وراء البيانات الوصفية والإدارية داخلياً، وبناء على ذلك يمكن استخدام ميتس مع المجموعات الرقمية التي تتفاوت حاجاتها لما وراء البيانات الوصفية تفاوتاً كبيراً، كما يمكن عن طريقه دمج ما وراء البيانات الفنية لأي صيغة من صيغ الملفات طالما أن مجتمعاً معيناً قد حدد خطة لغة ترميز موسعة تصف تلك الصيغة .

وكما أشرنا من قبل ، يمكن الإحالة إلى ما وراء البيانات الوصفية في خطة ميتس إما عن طريق المؤشر pointer ، أو عن طريق تضمين هذه البيانات داخل أحد العناصر المغلفة المسمى (<mdWrap>) ، وكخيار ثالث يمكن معاملة ما وراء البيانات الوصفية التي ترد في غير صيغة لغة الترميز الموسعة non-XML ، كسلسلة متتابعة من البيانات المستقلة ؛ ومن ثم تسجيلها داخل العنصر الفرعي المسمى (<binData>) من داخل العنصر المغلف المشار إليه من قبل، وهو (<mdWrap>) ، وهذا ما سيمكّن معيار ميتس من حمل أو استيعاب تسجيلة متضمنة في صيغة مارك . واستجابة لذلك يحتوي الموقع الرسمي لهذا المعيار على الوب على خطط توسعة كثيرة لما وراء البيانات التي يمكن تضمينها داخل ميتس، منها خطتا مودس MODS (إحدى الخطط التي طورتها مكتبة الكونجرس في صيغة لغة الترميز الموسعة لتتضمن الدلالات الخاصة بصيغة مارك)، ودبلن كور، كما أن هناك خطة توسعة أخرى تعرف بخطة ما وراء البيانات الوصفية

النوعية Generic Descriptive Metadata(GDM)، وهي خطة تتضمن عناصر ما وراء البيانات التي كانت متضمنة في معرف نوع الوثيقة الخاص بمعيار نهضة أمريكا الثاني القديم MOA٢ DTD ، والتي عزلت عن هذا المعيار وعرفت كخطة مستقلة بذاتها .

وفي سياق معيار ميتس خصصت أربعة أنواع لعناصر ما وراء البيانات : عنصر خاص بما وراء البيانات الفنية ويسمى <techMD>، وعنصر خاص بحقوق الملكية الفكرية ويسمى <rightsMD>، وعنصر خاص ببيانات المصدر الرقمي أو التناظري الذي أخذ عنه الكيان محل الوصف ويسمى <sourceMD>، وعنصر خاص بأصل الكيان الرقمي ويسمى <digiprovMD> . وفي داخل كل عنصر من العناصر يشار إلى ما وراء البيانات نفسها إما عن طريق الموجه إلى التسجيلة الخارجية، أو بتضمينها داخل وثيقة ميتس في داخل العنصر المغلف لما وراء البيانات المسمى <mdWrap>، ويلاحظ في هذا الصدد أن خطة ما وراء البيانات الفنية للصور الرقمية الثابتة التي أعدتها المنظمة الوطنية لمعايير المعلومات في صيغة لغة الترميز الموسعة المعروفة بخطة مكس MIX اعتبرت كخطة توسعة لما وراء البيانات الفنية .

وبحكم اعتماد معيار ميتس على خطط ما وراء البيانات الإدارية والفنية المعرفة خارجياً ، ينظر إليه على أنه من معايير ما وراء البيانات البنائية ، وهذا ما يبدو مجسداً في قسمي " مجموعة الملف " file group ، " والخريطة البنائية " structural map. حيث يجمع أول هذين القسمين ، أي قسم " مجموعة الملف (<fileSec>) " - وكما هو الحال في معيار نهضة أمريكا

الثاني المعلومات عن الملفات ذات العلاقة بنسخة رقمية واحدة من المصدر ، داخل عنصر التغليف المسمى <fileGrp> . وفي المثال التالي لدينا مجموعتان للملفات ، إحداهما للنسخة التي جاءت في صيغة تيف وهي لنشرة من صفحتين ، والأخرى للنسخة التي جاءت في صيغة جي بيج JPEG من النشرة نفسها. ومن الملاحظ أن الخاصية التي تعطى لعنصر الملف <file>، والمسماة "groupid" ، هي التي تستخدم لربط الملفات ذات الصلة بعضها ببعض، حيث نجدها في هذا المثال تربط بين النسختين المختلفتين من حيث الصيغة، وهما : QA٠٠٠.TIFF، و QB٠٠٠.JPG، عن طريق إعطائها القيمة نفسها التي تعطي لعنصر المجموعة وهي " groupid " . وقد جاء ذلك على النحو التالي:

```
<fileGrp>
    <file MIMEYPE="image/tiff" ID="QA٠٠٠" SEQ="١"
    GROUPID="١" ADMID="A١">
        <FLocat LOCTYPE="OTHER"
        OTHERLOCTYPE="PATH"
            xlink:type="simple"
            xlink:href="/sun٦/texts/QA٠٠١.TIFF"/>
    </file>
    <file MIMETYPE="image/tiff" ID="QA٠٠١"SEQ="٢"
    GROUPID="٢" ADMID="A١">
        <FLocat LOCTYPE="OTHER"  OTHERLOCTYPE="PATH"
            xlink:type="simple"
            xlink:href="/sun٦/texts/QA٠٠٢.TIFF"/>
    </file>
</fileGrp>
```

```
<fileGrp>
        <file MIMEYPE="image/jpeg" ID="QB٠٠٠" SEQ="١"
        GROUPID="١" ADMID="A٢">
                <FLocat LOCTYPE="OTHER"  OTHERLOCTYPE="PATH"
                        xlink:type="simple"
                        xlink:href="/sun٦/texts/QB٠٠١.JBEG"/>
        </file>
        <file MIMETYPE="image/jpeg" ID="QB٠٠١"SEQ="٢"
        GROUPID="٢" ADMID="A٢">
                <FLocat LOCTYPE="OTHER"  OTHERLOCTYPE="PATH"
                        xlink:type="simple"
                        xlink:href="/sun٦/texts/QB٠٠٢.JPEG"/>
        </file>
</fileGrp>
```

ومن الملاحظ كذلك أن عنصر الملف <file> له خصائص أخرى اختيارية، منها : خاصية نوع ملف المايم MIME ، وخاصية حجم الملف بالبايتات ، وخاصية تاريخ الإنشاء، وخاصية مجموع التدقيق checksum. كما تستخدم الخاصية "admid" لربط معلومات الملف الموضحة في قسم ما وراء البيانات الإدارية بالمعلومات [في هذا العنصر]، كما يستخدم العنصر الفرعي المسمى <Flocat> للربط بالملف الفعلي.

أما قسم الخريطة البنائية <structMap> ، والذي يعرف عادة " بقلب ميتس the heart of METS " ، فهو القسم الإجباري في هذه الخطة، وهو يعنى بتوضيح البنية الهرمية للأقسام المنطقية في داخل الكيان [الموصوف]، ثم ربطها بالملفات المختلفة[التي يتكون منها]. وفي الفقرة

التالية مثال لوثيقة ذات بنية داخلية محدودة ، وهي عبارة عن نشرة مكونة من صفحتين .

```
<structMap TYPE="logical">
        <div LABEL="Final Report of the Committee"
        TYPE="pamphlet">
                <div LABEL="Page ١" ORDER="١"
                ORDERLABEL="١"
                TYPE="page">
                        <fptr FILEID="QA٠٠٠"/>
                        <fptr FILEID="QB٠٠٠"/>
        </div>
                <div LABEL="Page ٢" ORDER="٢" ORDERLABEL="٢"
                TYPE="page">
                        <fptr FILEID="QA٠٠١"/>
                        <fptr FILEID="QB٠٠١"/>
                </div>
        </div>
</structMap>
```

ومن الملاحظ أن القيمة(١) التي أعطيت للخاصية المسماة " وسيمة label " في العنصر عالي المستوى المسمى <div>، يقصد بها أن يعرض هذا العنوان في قائمة المحتويات التي تعد لهذه الوثيقة، أما الصفحتان اللتان تتكون منهما هذه النشرة فيتم تمثيلها عن طريق العنصرين <div>

(١) يقصد بالقيمة هنا عنوان الوثيقة وهو" Final Report of the Committee " [المترجم].

المدرجين تحت العنصر الأكبر المسمى <div> الممثل للنشرة كلها، أما الخاصية المسماة " الترتيب Order " فتشير إلى ترتيب الصفحات في الوثيقة الأصلية، كما تشير القيمة l " وسيمة الترتيب Orderlabe " إلى الشكل الذي يظهر به رقم الصفحة في الوثيقة ، كذلك تشير الخاصية " الوسيمة lable " إلى الطريقة التي يجب أن يعرض بها رقم الصفحة للمستفيد ، أما بالنسبة للملفين الممثلين لهاتين الصفحتين فقد عرفا عن طريق القيمة التي أعطيت للخاصية " تعريف الملفات fileid " للعنصر المسمى " <fptr> " الذي يعنى بربط الصفحات بالملفات المعرفة في قسم "مجموعة الملف "وذلك عن طريق الخاصية المسماة id "التعريف" لعنصر الملف <file> .

وفضلاً عن ذلك ، يتوافر لميتس إمكانية الإشارة إلى أجزاء الملفات كذلك؛ حيث يمكن إضافة العنصر الفرعي <area> داخل العنصر <fpt> ، ومن خلاله يتم الربط بين الأقسام الفرعية للملف . كذلك هناك الخاصيتان " البداية Begin "، و" النهاية end " اللتان تحددان موضعي البداية والنهاية داخل الملف المشار إليه، كذلك هناك الخاصية "betype" التي تبين تحديد نقطتي البداية والنهاية، كأن يشار فيها - على سبيل المثال - إلى ذلك هكذا : as byte offsets ، أو MIDI ، أو SMIL (١).

ولعل ما تجدر الإشارة إليه في هذا الصدد أن لخطة ميتس أوجه استخدام كثيرة متوقعة ، إذ يمكن استخدامها في نقل الكيانات الرقمية عبر

---

(١) تشير القيمة الثانية إلى رمز الوقت المخصص لعرض المواد الموسيقية، في حين تشير الثالثة إلى الوقت المخصص لتقديم العروض السمعية والبصرية [المترجم].

المستودعات ، وفي ضبط عمليتي عرض المصادر وتصفح المستفيد النهائي لها ، ونتيجة للقدرة التي تتميز بها ميتس كذلك على وضع encapsulate كل من ما وراء البيانات الوصفية والبنائية والإدارية الخاصة بمصدر المعلومات الموصوف في تركيبة واحدة، يمكن كذلك استخدامها " كحزمة تقديم معلومات (Submission Information Package (SIP" للأرشيفات الرقمية التي تتبع إطار نظام المعلومات الأرشيفة المفتوح OAIS . وتأكيداً لذلك فقد استخدمت جامعة هارفارد ، كجزء من دراسة تقدمت بها مؤسسة أندرو ميلون لأرشفة الدوريات الإلكترونية ، خطة ميتس كأساس لخدمة تقديم المعلومات التي سيستخدمها الناشرون للمشاركة بإصدارات دورياتهم الإلكترونية في مستودع الجامعة .

إن صيغة ميتس لهي بالفعل صيغة مرنة إلى أبعد حد ، وبدرجة تسمح بتنفيذها بطرق متعددة ، كما أنها تعد خطة قابلة للتوسع بلا حدود ، نتيجة اعتمادها على خطة التوسعة ، ولهذا السبب تلقى توجهات تطوير بروفايلات تطبيق ميتس التي توثق استخدام عناصرها ، وتوجهات تطوير خطة خارجية لتطبيقات معينة لميتس دعماً وتشجيعاً على استمرارها ، وتبعاً لذلك تقوم مكتبة الكونجرس بصفتها الجهة المخولة برعاية هذه الخطة بتطوير نظام تسجيل مركزي لبروفايلات تطبيق ميتس .

صيغة إم بيج -٧ **MPEG-٧** :

كما هو الحال مع خطة ميتس ، تقوم صيغة بتغليف ما وراء البيانات الوصفية والإدارية والبنائية(١). وقد تبنى جماعة خبراء الصور المتحركة

(١) أي بوضعها في تركيبة واحدة [المترجم].

التابعة للمنظمة الدولية للتقييس واللجنة الدولية للتقنيات الإلكترونية ISO/IEC تطوير هذه الصيغة ، والتي تركز بشكل أساسي على المحتوى السمعي والبصري . وفي عام ٢٠٠١م أطلقت مواصفة هذه الصيغة، المواصفة رقم ١٥٩٣٨ ISO/IEC ، وقد جاءت في سبعة أجزاء[١] على النحو التالي:

- الجزء الأول : النظم .
- الجزء الثاني : لغة وصف التعريف Description definition language .
- الجزء الثالث : المواد البصرية .
- الجزء الرابع : المواد السمعية .
- الجزء الخامس : خطط وصف الوسائط المتعددة .
- الجزء السادس : البرمجية المرجعية reference software .
- الجزء السابع : اختبار التوافق conformance testing .

ومن الواضح أن الجزء الأول يعنى بعملية ترميز ونقل ما وراء البيانات في صيغ نصية وثنائية بلغة الترميز الموسعة في بيئة ديناميكية قد ترسل خلالها أجزاء الوصف بشكل غير منتظم ، كما يمكن أن ترسل خلالها - حسب الطلب - تحديثات كاملة أو جزئية للوصف (مثال ذلك : أضف /

(٢) تجدر الإشارة إلى أن كل جزء من هذه الأجزاء يحمل رقم المواصفة ثم رقم الجزء متبوعين بالجزء هكذا ISO/IEC ١٥٩٣٨-١: systems وهكذا [أي بوضعها في تركيبة واحدة] [المترجم].

احذف / استبدل) .

أما الجزء الثاني فيعرِّف بلغة وصف التعريف ، بهدف تحديد كل من الواصفات وخطط الوصف التي تعد بمثابة الوحدات الأساسية لوصف صيغة إم بيج -٧ . وفي هذا الصدد ينظر إلى الواصفة على أنها عنصر تعريف، وإلى خطة الوصف على أنها عناصر تعريف متصلة بعضها مع بعض، وعليه تعد لغة وصف التعريف بمثابة نسخة من تعريف خطة لغة الترميز الموسعة مصحوبة ببعض التوسعات لإضافة بعض الخصائص التي يحتاج إليها مجتمع المواد السمعية والبصرية .

أما الأجزاء من الثالث إلى الخامس فتعنى بتعريف الواصفات وخطط الوصف الخاصة بالمواد البصرية، ثم السمعية، ثم الوسائط المتعددة، على هذا الترتيب، بحيث إن كل جزء منها يتكون من عناصر ما وراء البيانات الفنية التي تتفق وطبيعة هذا النوع من المواد، إذ يشتمل الجزء الخاص بالمواد البصرية - على سبيل المثال - على خمس وعشرين واصفة أو خطة وصف تستخدم لوصف مقاطع الفيديو، والأجزاء المتحركة، والأجزاء الثابتة، وفي هذا الصدد تعرف عناصر ما وراء البيانات هذه المواد البصرية من خلال الخصائص التالية : اللون، القوام، الشكل، الحركة، الموضع، سمات الوجوه البشرية .

ويحدونا الأمل أن ينتهي الأمر بهذا المعيار (MPEG-٧) إلى أن يندمج في إطار العمل الخاص بالوسائط المتعددة ، المسمى " MPEG-٢١ "، والذي لا يزال قيد التطوير، ذلك أن معيار إم بيج - ٧ يهدف إلى وصف المصادر

والمساعدة على إدارتها، بينما سيقوم المعيار إم بيج ٢١ بوصف المصادر والمساعدة على إدارة تفاعل المستفيد معها، بما في ذلك إدارة الحقوق المرتبطة بذلك، ولذلك ينظر إلى معيار إم بيج - ١٢ على أنه مجموعة من المواصفات المترابطة بعضها مع بعض.

وأخيرًا تجدر الإشارة إلى أن الجزء الأول من هذا المعيار والذي يعكس رؤية هذا الإطار أو الأسس العامة له نشر في شكل تقرير فني للمنظمة الدولية للتقييس واللجنة الدولية للتقنيات الإلكترونية(١)، كما أن مسودة الأجزاء الأخرى يتم تداولها بين عدد من الجهات المعنية . وقد خطط أن تأتي على النحو التالي :

- الجزء الثاني : بيان المادة الرقمية .
- الجزء الثالث : هوية المادة الرقمية .
- الجزء الرابع : إدارة الملكية الفكرية وحمايتها .
- الجزء الخامس : لغة التعبير عن الحقوق .
- الجزء السادس : معجم بيانات الحقوق .
- الجزء السابع : تكييف المادة الرقمية .
- الجزء الثامن : البرمجية المرجعية .

READINGS :

---

(١) ISO/IEC TR ٢١٠٠٠-١:٢٠٠١ Information Technology-Multimedia Framework (MPEG- ٢١), Part ١: Vision, Technologies, and Strategy. International Organisation for Standardization (٢٠٠١).

- Digital Page Imaging and SGML: An Introduction to the Electronic Binding DTD (Ebind) (home page). Available at http://sunsite.berkeley.edu/Ebind/. Accessed ٢٢ July ٢٠٠٢.

  The Ebind home page has links to the downloadable DTD, sample documents in their encoded and viewable forms, and tools for Ebind use.

- METS-Metadata Encoding and Transmission Standard: Official Web Site (home page). Available at http://www.loc.gov/standards/mets/. Accessed ٢٢ July ٢٠٠٢.

  Links to the METS Schema, Extension Schema, and toolkits, as well as to news and announcements about METS, examples of METS-encoded documents, and a tutorial. The site will ultimately link to registries of METS application profiles and types.

- The MPEG standards, being ISO standards, must be purchased from ISO. However, there is ample MPEG information freely available online:

- Hunter, Jane. "MPEG-٧: Behind the Scenes:' D-Lib Magazine ٥, no. ٩ (September ١٩٩٩). Available at http://www.dlib.orgtdlib/september٩٩/hunter/٠٩hunter.html.

  Written while the standard was still under development, this article explains the objectives, uses, and components of MPEG- ٧.

- The MPEG Home Page. Available at http://mpeg.telecomitalialab.com/ . Accessed ٢٢ July ٢٠٠٢.

This has links to information about all the MPEG standards, including MPEG-٧ and MPEG-٢١. See particularly, in the Documents section, "MPEG-٧ Overview" (ISO/IEC JTC١/SC٢٩/WGII N٤٦٧٤).

- MPEG-٧ Home Page. Available at http://www.mpeg-industry.com/ . Accessed ٢٢ July ٢٠٠٢.

  This site is particularly useful for its tutorials on various parts of the MPEG- ٧ family of standards.

الفصل الثامن عشر

ما وراء البيانات الخاصة بالحقوق

لقد أثارت السهولة التي تحيط بعمليات إتاحة الوصول للمحتوى الرقمي واستنساخه وتعديله اهتماماً محموماً بإدارة الحقوق الرقمية، وقد برزت في هذا السياق عبارة " إدارة الحقوق الرقمية Digital Rights Management "، وهي عبارة يحيطها شيء من الغموض، كما أنها تستخدم للتعبير عن أمرين : أولهما إدارة الحاسب لحقوق الملكية الفكرية، وثانيهما إدارة الحقوق في المحتوى الرقمي the Management of Rights in Digital content (١). غير أن هناك تعريفاً أوسع من ذلك يتداول في عالمي المكتبات والنشر، وهو ما يتجسد فيما تراه مؤسسة معرف الكيان الرقمي الدولية International DOI Foundation (IDF) من أن النظام العملي لإدارة الحقوق ينبغي أن يغطي الإدارة الرقمية لجميع الحقوق الرقمية وغير الرقمية(٢)، ومع ذلك فيلاحظ أن هذه العبارة أي "إدارة الحقوق الرقمية " وخاصة عندما تكتب بالحروف الاستهلالية الكبيرة - هكذا Digital Rights Management - أو مختصرة - هكذا DRM - غالبًا ما ينحصر معناها في الغالب لتشير إلى تعزيز حماية المحتوى عن طريق إحدى البرمجيات .

ومن منطلق القناعة بأن ما وراء البيانات باتت مكوناً أساسياً من مكونات أي من نظم إدارة الحقوق، فقد استعرت جذوة الأنشطة الساعية إلى تطوير ما وراء بيانات خاصة بالحقوق لتستخدم في سياقات مختلفة. وسنلقي

(١) أو إدارة حقوق المحتوى الرقمي[المترجم].

(٢) Norman. Paskin, "Position Paper for W٣C Workshop on Digital Rights Management for the Web" (٢٢-٢٣ January ٢٠٠٠), available at

http://www.doi.org/٠٠١٢١٩W٣C.pdf. Accessed ٣١ July ٢٠٠٢.

الضوء في الفقرات التالية على بعض هذه الأنشطة التي يجري تطويرها الآن على الساحة.

مشروع إنديكس **<INDECS>** :

تكفل مشروع التشغيل البيني للبيانات في نظم التجارة الإلكترونية Interoperability of Data in E-Commerce Systems أو ما يعرف اختصاراً بـ " إنديكس INDECS ، والذي مولته المفوضية الأوربية خلال الفترة من ١٩٩٨ إلى ٢٠٠٠م، بتقديم رؤية عالية المستوى ومؤثرة لمتطلبات ما وراء البيانات اللازمة لإدارة الحقوق . وقد اتخذ هذا المشروع منذ بدايته منظوراً يتسم بأنه متعدد الجنسيات من ناحية، ومتعدد الوسائط من جهة أخرى ، مما جعله يحظى بدعم دولي من المؤسسات التجارية الكبرى المتمثلة في شركات إنتاج الصوتيات ، وناشري المواد الموسيقية، وشركات إنتاج المرئيات ، وناشري الكتب والدوريات . وقد تجسدت أهداف هذا المشروع في إنشاء إطار للتجارة الإلكترونية لحقوق الملكية الفكرية لجميع الوسائط ، وكان من باكورة جهود هذا المشروع الوثيقة التي نشرت في صيف عام ٢٠٠٠م بعنوان " إطار إنديكس لما وراء البيانات : المبادئ، والنموذج ، ومعجم البيانات :

the <indecs> Metadata framework: Principles, Model and Data Dictionary " (http://www.indecs.org/pdf/framework.pdf)

ويعد نموذج إنديكس بمثابة نموذج دلالي لوصف الملكية الفكرية من جهة، والأطراف التي ترتبط بإنشائها وتداولها تجارياً من جهة ثانية ،

والاتفاقيات التـي تبـرم حـولها مـن جهـة ثالـثة. ويـنطلق هـذا المـشروع مـن المسلمات التـي تـرى أن خططـاً مختلفة لمـا وراء البيانات سيتم تطويـرها ، وأن هذه الخطـط ستستخدمها قطاعـات صـناعية معينـة (مثل ناشري الكتب والمـواد الموسيقية)، ودفعاً لمسيرة التجارة الإلكترونية العالمية ، ينبغي ألا تكـون هناك معـوقات لتبادل مـا وراء البيانات هذه من جهة ، وأن يكون من الممكن إعادة استخدامها في سياقات مختلفة من جهة أخرى .

ويحـاول مـشروع إنـديكس تجمـيع العناصـر الوصـفية الأساسـية التـي لا يمكـن الاسـتغناء عنهـا عند وصـف الحقـوق فـي مجموعة محـددة من الفئات والقيم النوعية القابلـة للتطبـيق علـى المستوى العالمـي ، ذلك انطلاقاً من أن البيـانات يمكـن تـبادلها بـين خطـط مـا وراء البيانات المتخصصة في مجالات معينـة إذا مـا أخـذت عناصـر البيـانات مـن معجـم البيانات الخـاص بمشروع إنديكس <indecs> data dictionary ، أو كان بالإمكان ربط هذه العناصر بهذا المعجـم ؛ حـيث يمكـن - علـى سـبيل المـثال - أن يكـون فـي إحـدى خطـط مـا وراء البـيانات عنـصـر خـاص "بالمـشارك contributor " ، وقـيمة لـدور " مخـرج الـنص المسرحي screenplay adapter " ، بينما يكون في خطـة أخرى عنـصر " للمـوزع الموسـيقي musical arranger " . وبتـرجمة هذه الوظائف إلـى مـصطلحات إنـديكس ، نجـدها بمثابة نماذج محددة للفئة النوعية (الدور العاملـي للمـشارك contributor agent role) ، وللقـيمة (القـائم بالـتعديل modifier).

ومـع أن إطـار إنـديكس لا يقوم بتحليل أية خطة محددة ومعرفة خارجياً مـن خطـط ما وراء بيانات ، إلا أنه يضع أربعة مبادئ إرشادية لتطوير " ما

وراء بيانات جيدة الصياغة well formed " تهدف إلى دعم تجارة إلكترونية فاعلة " . وينص أول هذه المبادئ - والذي يعرف " بمبدأ الهوية الفريدة the principle of unique identification " على وجوب تعريف كل كيان تعريفاً فريداً داخل اسم مميز namespace . ونظراً لأن الكيان يعرف على أنه " شيء ما تم التثبت من هويته something which is identified " ، فتجدنا أمام دائرة من الجدليات التي لا تنتهي [حول هذا الموضوع] ، مع أن الفكرة الأساسية واضحة ، ومفادها أن المعرفات الفريدة ينبغي أن تستخدم للأجزاء وللأشياء ، كما ينبغي أن تؤخذ قيم عناصر ما وراء البيانات الوصفية من لغات مقيدة معروفة .

ويدعو المبدأ الثاني - الذي يعرف " بمبدأ التفتيت الوظيفي the principle of functional granularity " - إلى أن تكون ما وراء البيانات كفيلة بتعريف أجزاء مصادر المعلومات بإصداراتها المختلفة عند أي مستوى من التفصيل الدقيق، مادامت هناك حاجة إلى هذا التفصيل . أما المبدأ الثالث - الذي يعرف "بمبدأ المرجعية المخصصة the principle of designated authority " - فيشترط أن تحقق هوية أي من مؤلفي مواد ما وراء البيانات بطريقة محددة يمكن معها التثبت من هويته ، وأخيراً يأتي المبدأ الرابع - الذي يعرف بمبدأ "إتاحة الوصول المناسب the principle of appropriate access " - ويشير إلى أن ما وراء البيانات ينبغي أن يتاح الوصول إليها عندما يتطلب الأمر ذلك، وينبغي أن تكون في الوقت ذاته مؤمنة من انتهاكات الاستخدام غير المشروع .

لقد كان لإطار إنديكس أثره الواضح في تشكيل الرؤى حول ما وراء

البيانات الخاصة بالحقوق بدرجة دفعت كلاً من إديتير EDItEUR (التي تعد المنظمة الأم الراعية لمجموعة معايير أونكس) ، والمؤسسة الدولية لمعرف الكيان الرقمي IDF إلى دعم مسيرته، بل الأكثر من ذلك أن هذا الإطار اتُخذ أساساً لمبادرة إنديكس آر آر دي <indecs>rrd ، تلك المبادرة ذات الطابع الائتلافي التي تهدف إلى تطوير معجم بيانات لما وراء بيانات الحقوق ، وفي عام ٢٠٠١م اختيرت مواصفة تصميم هذه المبادرة - أي إنديكس آر آر - لتكون أساسًا يبنى عليه الجزء السادس المخصص لمعجم بيانات الحقوق من معيار أم بيج - ١٢ MPEG-١٢ الذي سبقت الإشارة إليه في الفصل السابق ، والذي يجري العمل على تطويره الآن، ويتوقع أن يقر في ٢٠٠٣م.

الكتاب الإلكتروني المفتوح :

عرف عن منتدى الكتاب الإلكتروني المفتوح[١] (OeBF) The Open eBook Forum بأنه الجهة التي تبنت تطوير تركيبة نشر الكتاب الإلكتروني المفتوح OEB Publication Structure، وهي إحدى مواصفات الترميز المعياري لمحتوى الكتاب الإلكتروني ، وفضلاً عن ذلك يقوم هذا المنتدى بعدة أنشطة أخرى ، منها اضطلاعه بمهام مجموعة العمل الخاصة بالحقوق والقواعد Rights and Rules Working Group المكلفة بتطوير معيار مفتوح وقابل للتطبيق على نحو تجاري لضبط عمليات التشغيل البيني لنظم إدارة الحقوق الرقمية DRM، ومن ثم إتاحة إمكانات التبادل الآمن للمطبوعات الإلكترونية (ePublications) فيما بين الأطراف المرتبطة بها والمتمثلة في

(١) لغة يقال في الإنجليزية للكتاب المفتوح open letter ، وقد استخدمناها مجازًا في هذا السياق ، مع أنها قد تشير إلى المواصفة قيد الدراسة والبحث [المترجم].

أصحاب الحقوق، والوسطاء، والمستفيدين(١).

لقد انبثقت [مجموعة] الحقوق والقواعد بالفعل من جراء اندماج كل من منتدى الكتاب الإلكتروني المفتوح ومجموعة العمل الخاصة بتبادل الكتاب الإلكتروني (Electronic Book eXchange(EBX)(٢)، تلك اللجنة التي كانت تعمل منذ عام ١٩٩٨م على تطوير معايير لصيغة محتوى الكتاب الإلكتروني من جهة، وحماية حقوق ملكيته وتوزيعه من جهة أخرى . و في عام ٢٠٠٠م أطلقت مواصفة نظام تبادل الكتاب الإلكتروني Electronic Book Exchange System (EBX) ، مشتملة على وصف للكيفية التي تتفاعل بها المكونات المؤمنة أو الموثوق فيها trusted في نظام معين يحمي الملكية الفكرية عبر عمليات التفاعل بين ناشري هذه الكتب ، وبائعيها وموزعيها ، والمكتبات التي تقتنيها ، والمستفيدين الذين يستخدمونها .

وتتمثل الآلية المستخدمة لتفعيل ذلك عبر نظام تبادل الكتاب الإلكتروني في إنشاء وتداول كيانات رقمية تسمى " مستندات القيد أو الإثباتات vouchers "، وهي عبارة عن وصف مرمز في صيغة لغة الترميز الموسعة للتصاريح التي تصاحب ملف الكتاب الإلكتروني ، والتي تختلف باختلاف المراحل التي يمر بها الكتاب الإلكتروني عبر رحلة تداوله ، ومن بين التصاريح يمكن أن يبين مستند قيد الكتاب الإلكتروني ما إذا كان هذا الكتاب (أو

---

(٢) Open eBook Forum Rights and Rules Working Group (web page), available at http://www.openebook.org/members/Rights-Rules/index.htm. Accessed ٣١ July ٢٠٠٢.

(٣) يلاحظ هنا أن المختصر الدال على كل من نشاط تبادل الكتاب الإلكتروني الذي تبنته لجنة العمل هو نفسه الدال على نظام تبادل الكتاب الإلكتروني ، غير أن الأول يكتب هكذا Electronic Book eXchange(EBX) ، بينما يكتب الثاني Electronic Book Exchange System (EBX) [المترجم].

بالأحرى مستند القيد) متاحًا للإعارة ، أم للإهداء ، أم للبيع ، كما يمكنه أن يبين كذلك الفترة التي يسمح فيها لصاحب ملكية الكتاب أن يعير مستند القيد ، وأن يبين كذلك بعض الضوابط الخاصة بالاستخدام الشخصي للكتاب مثل الحد الأقصى لعدد النسخ المسموح بها ، وطول المدة ، ومقدار المحتوى المسموح به كذلك . وفضلاً عن ذلك تعرف المواصفة طبيعة التفاعل بين نظام القراءة وخادم مستند القيد voucher server ، كما تعمل على تعريف إحدى اللغات الشاملة لإدارة الحقوق .

ونتيجة للرعاية التي أحاطت بكل من منتدى الكتاب الإلكتروني المفتوح ومواصفة تبادل الكتاب الإلكتروني من قبل الجهات نفسها، مع تداخل مصالحها في ذلك ، فقد توحدت جهود الفريقين في خريف عام ٢٠٠٠م، وأدمجت مبادرة تبادل الكتاب الإلكتروني المفتوح في المنتدى ، واستمرت أعمالها من خلال مجموعة عمل المنتدى (OeBF) الخاصة بالحقوق والقواعد the OeBF Rights and Rules Working Group ، وتعكف هذه المجموعة في الوقت الحاضر على التوصل إلى مجموعة من متطلبات أو قواعد الحقوق rights grammar ، بحيث تكون شاملة لكل من لغة التعبير عن الحقوق rights expression language ، ومعجم بيانات الحقوق rights data dictionary.

وتجدر الإشارة في هذا الصدد إلى أنه في الوقت نفسه تقريباً الذي صدرت فيه مواصفة تبادل الكتاب الإلكتروني ، أصدرت جمعية الناشرين الأمريكيين مواصفتها الخاصة بمتطلبات الناشرين لإدارة الحقوق الرقمية

للكتب الإلكترونية[1]، وتحدد هذه المواصفة التي كتبها أحد الاستشاريين المتعاقدين مع جمعية الناشرين الأمريكيين كلاً من أساليب التسعير pricing وسيناريوهات الاستخدام التي ينبغي أن تدعمها لغة إدارة الحقوق، كما تدعو الناشرين للمشاركة في إحدى مجموعات معايير الكتاب الإلكتروني المفتوحة، مثل منتدى الكتاب الإلكتروني المفتوح OeBF.

لغة الحقوق الرقمية المفتوحةODRL ، واللغة الموسعة لترميز الحقوق XRML :

يتكون النموذج العام لإدارة الحقوق من ثلاثة مكونات رئيسية هي : العملاء (أفراداً وهيئات)، والملكية الفكرية، والاتفاقيات التي تحكم طبيعة العلاقة بين العملاء والملكية الفكرية. وقد عبر إطار إنديكس عن ذلك بعبارة بليغة تقول: "يعقد الناس صفقات حول الأشياء people do deals about stuff "[2]. ولطالما تعد المعلومات الببليوجرافية التي تصف الملكية الفكرية من قبيل ما وراء بيانات، فإن المعلومات التي تصف المعاملات أو الاتفاقيات التي تدور حول الحقوق تعد تبعًا لذلك من قبيل ما وراء بيانات كذلك ، لكن ليس من الشائع أن تسمى الأساليب أو الوسائط التي تتداول بها

---

(١) Association of American Publishers, Digital Rights Management for Ebooks: Publisher Requirements, version ١,٠ (٢٠٠٠), available at

http://www.publishers.org/drm.pdf. Accessed ٣١ July ٢٠٠٢.

(١) Godfrey Rust and Mark Bide, The <indecs> Metadata Framework: Principles, Model and Data Dictionary (June ٢٠٠٠), p. ٤, available at

http://www.indecs.org/pdf/framework.pdf. Accessed ٣١ July ٢٠٠٢.

المعلومات حول الحقوق خطط ما وراء بيانات خاصة بالحقوق ، وإنما تسمى نظم أو لغات حقوق rights languages ، وبشكل رسمي تعرف هذه اللغات إما بلغات التعبير عن الحقوق الرقمية Digital Rights Expression Languages(DREL) ، أو بلغات الحقوق الرقمية Digital Rights Languages(DRL) . ومن أشهر النظم أو اللغات السائدة في هذا الصدد: لغة الحقوق الرقمية المفتوحة the Open Digital Rights Languages(ODRL) ، واللغة الموسعة لترميز الحقوق the eXtensible rights Markup Language(XrML) .

وقد تبنت نظم حقوق الملكية الفكرية في أستراليا تطوير أولى هاتين اللغتين - أي لغة الحقوق الرقمية المفتوحة (ODRL) - ويروج لها الآن على أنها أحد المعايير المفتوحة، وفي سياق نموذج هذه اللغة يصطلح على تسمية العملاء "بالأطراف parties "، وعلى الملكية الفكرية " بالأصول أو الممتلكات assets" ، وتعتمد هذه اللغة في تعريفها الفريد لكل من الأطراف والممتلكات على خطط خارجية ، كما أنها تركز في هذا الصدد على عملية إقرار الحقوق أو التعبير عنها، والتي رأت أن تمثلها تسعة كيانات هي: التصاريح ، والقيود، والمتطلبات، والشروط أو الضوابط، وأصحاب الحقوق ، والسياقات ، والعروض ، والاتفاقيات، وإلغاء الاتفاقيات revocation . مع ملاحظة أن التصاريح تعنى بضبط إجراءات استخدام الممتلكات أو الأصول وإعادة استخدامها ونقلها وإدارتها ؛ وبالتالي يمكن - على سبيل المثال - أن يمنح التصريح باستخدام أحد الممتلكات عن طريق عرضها وطباعتها . أما القيود فتعنى بتقييد طبيعة هذه التصاريح ؛ حيث يمكن - على سبيل المثال -

أن تحدد القيود التي تفرض على المستفيدين أن تصاريح الاستخدام التي منحت لهم مكفولة لهم وحدهم أو لجماعة محددة وليس لدون ذلك . أما المتطلبات فهي بمثابة شروط مسبقة ينبغي أن تتوافر للحصول على التصاريح ؛ كأن يشترط الدفع المسبق للرسوم . وأما الشروط فهي بمثابة الضوابط التي ينبغي الالتزام بها ، وما لم تراع هذه الضوابط فستنتهي فعالية تصاريح الاستخدام.

وتجدر الإشارة إلى أن هذه التصاريح والقيود والمتطلبات والشروط وغيرها من الكيانات الأخرى يمكن أن تمثل في صيغة لغة الترميز الموسعة ، وعليه فإن مواصفة لغة الحقوق الرقمية المفتوحة ODRL تحتوي على النموذج الذي يعرف بالكيانات ومكوناتها ، كما تشتمل كذلك على معجم بيانات يعرف بدلالات جميع العناصر المستخدمة في لغة التعبير عن الحقوق الخاصة بلغة الحقوق الرقمية المفتوحة ، وكذلك على طريقة الترميز في صيغة لغة الترميز الموسعة لكل من مصطلحات لغة الحقوق الرقمية المفتوحة وعناصرها ، كما تحتوي هذه المواصفة أيضاً على قسم يبين الكيفية التي تعرف بها أية معاجم بيانات إضافية.

أما بالنسبة للغة الموسعة لترميز الحقوق eXtensible rights Markup Language(XrML) ، فهي إحدى منتجات شركة "كونتنت جارد CotentGuard " التي تملكتها شركتا زيروكس ومايكروسوفت ، وكانت هذه الشركة قد أعلنت في إبريل ٢٠٠٢م أنها جمدت أنشطتها الخاصة بتطوير لغة ترميز الحقوق الموسعة XrML عند الإصدار رقم ٢,٠ ، وأوكلت ذلك

إلى منظمة ترقية أو تطوير معايير المعلومات البنائية OASIS [(١)]، إحدى الائتلافات الدولية التي تركز جهودها على تطوير مواصفات معايير [صناعية] تقوم على صيغة لغة الترميز الموسعة. وبموجب ذلك شكلت لجنة فنية خاصة بلغة الحقوق تابعة لمنظمة تطوير معايير المعلومات البنائية OASIS Rights Language Technical Committee ، تضم ممثلين لشركات : "كونتنت جارد ContentGuard " ، و"هيوليت باكارد Hewlett-Packard" ، و"مايكروسوفت" ، و"رويترز Reuters " ، و"فيريساين VeriSign " وغيرها من الشركات الكبرى الضالعة في صناعة إدارة الحقوق الرقمية DRM ، بهدف ترقية لغة ترميز الحقوق الموسعة XrML لتصبح معياراً خاصًا بلغة الحقوق ، وفضلاً عن ذلك تجدر الإشارة إلى أن هذه المواصفة - أي لغة ترميز الحقوق الموسعة - كان قد ووفق على اتخاذها أساساً بني عليه القسم الخامس من معيار MPEG-٢١ والذي جاء باسم " لغة التعبير عن الحقوق" . وعليه فمن المرجح أن تطور هذه المواصفة لتصبح لغة التعبير عن الحقوق المهيمنة على صناعة المحتوى .

ومع أن لغة ترميز الحقوق الموسعة يحسب لها أنها أكثر نضجاً وشمولاً من لغة الحقوق الرقمية المفتوحة ODRL ، فمن الملاحظ أن كلتا

---

(١) المختصر يعبر عن Organization for the Advancement of Structured Information Standards وهي عبارة عن ائتلاف دولي غير ربحي تأسس عام ١٩٩٣م، يتبنى تطويرَ وتبني معايير الأعمال الإلكترونيةِ e-business standards . ويشارك في هذا الائتلاف الآن أكثر من ٥٠٠٠ عضو تمثل أكثر من ٦٠٠ جهة من حوالي ١٠٠ دولة [المترجم عن موقع هذا الائتلاف على الإنترنت].

المواصفتين تشتركان في تغطية مجال الاهتمام نفسه ومن المنظور نفسه أو الرؤية كذلك ؛ فكلتاهما تجسد وجهة نظر الناشر أو المنتج ، وليس وجهة نظر المؤلف أو المستفيد من المحتوى من ناحية ، كما أنهما تتناولان قضايا الاستخدام ، وليس حقوق إتاحة الوصول من ناحية ثانية ، كما أنهما تسلمان من ناحية ثالثة بأن ما لم يمنح صراحة من التصاريح يعد لاغياً ولا يعتد به . ولعل آخر ما تجدر الإشارة إليه في هذا الصدد أن المجتمعات البحثية والتعليمية تحاول - كردة فعل لهذا التوجه - تطوير إطار خاص بها لإدارة الحقوق الرقمية ، بحيث يعنى في جانب منه بحقوق إتاحة الوصول إلى المصادر المرخص باستخدامها ، ويعنى في جانبه الآخر بحقوق المستفيدين وأصحاب الحقوق على حد سواء ، وسيكون من أهم ما تسلم به هذه المبادرة العمل على تطوير مجموعة أساسية من ما وراء بيانات إدارة الحقوق لتستخدم في هذا السياق.

READINGS :

- Coyle, Karen. "Stakeholders and Standards in the E-book Ecology: Or, It's the Economics, Stupid!" Library Hi Tech ١٩, no. ٤ (٢٠٠١): ٣١٤-٣٢٤.

  This knowledgeable and insightful article covers digital rights management as well as other standards efforts related to e-books.

- Iannella, Renato. "Digital Rights Management (DRM) Architectures." D-Lib Magazine ٧, no. ٦ (June ٢٠٠١). Available at http://www.dlib.org/dlib/juneO١/iannella/٠٦iannella.html.

A clear overview of two architectures for digital rights management, the "functional" architecture and the "information" architecture, written by one of the principal developers of ODRL.

- Martin, Mairead, et al. "Federated Digital Rights Management: A Proposed DRM Solution for Research and Education." D-Lib Magazine ٨, no. ٧/٨ (July/August ٢٠٠٢). Available at http://www.dlib.org/dlib/julyO٢/martin/٠٧martin.html.

  A description of how some in the research networking and library communities are attempting to develop a digital rights management architecture for teaching and research that uses Internet٢ middleware.

قائمة المصطلحات والمختصرات

A-core

دبلن كور الإداري .

AAP (Association of American Publishers)

جمعية الناشرين الأمريكيين .

AAT (Art and Architecture Thesaurus)

مكنز الفنون والعمارة .

Actionable

رابط فعال أو قابل للتفعيل .

ADL (Advanced Distributed Learning)

التعلم الموزع أو اللامركزي المتطور : مبادرة لوزارة الدفاع الأمريكية .

Administrative metadata

ما وراء البيانات الإدارية

AIIM (International the Association for Information and Image Management)

الجمعية الدولية لإدارة المعلومات والصور.

AITF (the Art Information Task Force of the College Art Association and the Getty Art History Information program)

لجنة معلومات الفنون لجمعية كليات الفنون وبرنامج المعلومات لمعهد جيتي لتاريخ الفنون : هي اللجنة التي طورت فئات وصف أعمال الفنون.

ALA (American Library Association)

جمعية المكتبات الأمريكية .

AMC

(آ إم سي) مصطلح استخدم مع صيغة مارك الامريكي USMARC للضبط الأرشيفي والمخطوطات ، هذا المصطلح لم يعد مستخدمًا مع تكامل صيغة مارك الأمريكي USMARC Format Integration .

ANSI (American National Standards Institute).

معهد المعايير الوطنية الأمريكية .

Application profile see profile

بروفايل تطبيق

APPM (Archives, Personal Papers, and Manuscripts)

دليل فهرسة المستودعات الأرشيفية والجمعيات التاريخية ومكتبات المخطوطات.

Archival collection

المجموعة الأرشيفية .

Archives

الأرشيفات، أو المحفوظات .

ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe)

أريادين : اتحاد شبكات التأليف والنشر التعليمي عن بعد لأوربا .

ArtSTOR

آرت ستور : مشروع تبنته مؤسسة ميلون لتطوير وتخزين وتوزيع الصور الرقمية وما يتصل بها من مواد بحثية إلكترونياً ؛ بهدف خدمة أغراض دراسة تاريخ الفنون والعمارة وغيرهما من مجالات العلوم الإنسانية.

Attribute

خاصية .

Attribute set

مجموعة الخصائص .

Authority file

ملف استنادي .

BASIC(Book And Serials Industry Communication)

منتدى المعايير لمجموعة دراسة صناعة الكتب .

Bib-١

بيب - ١ :مجموعة صفات في معيار Z٣٩,٥٠ طورت أساساً لبحث تسجيلات مارك .

BIBLINK

بيب لينك : مشروع موله الاتحاد الأوربي لإيجاد علاقة بين الهيئات الببليوجرافية الوطنية وناشري المواد الإلكترونية.

Bibliographic utility

مرفق ببليوجرافي .

BIC (Book Industry Communication)

بيك : منظمة بريطانية تطور وتروج المعايير الخاصة بالتجارة الإلكترونية والاتصالات في صناعة الكتاب والمسلسلات .

BICI (Book item and contribution identifier)

معرف مكونات الكتاب وأجزائه .

BSR (the ISO Basic Semantics Register)

مسجل الدلالات الأساسية لمنظمة الأيزو .

CC:DA (Committee on Cataloging: Description and access)

لجنة تابعة للجمعية الأمريكية للمكتبات مهمتها متابعة قواعد الفهرسة بالتعديلات ، وتسمى لجنة الفهرسة : الوصف والوصول.

CDF (Channel Definition Format)

صيغة تعريف القناة .

CDWA (Categories for the Description of the Works of Arts)

فئات وصف الأعمال .

CEDARS (CURL Exemplars in digital Archives)

سيدارز : مشروع بريطاني لاستكشاف قضايا الحفظ الرقمي .

CEMARC (curriculum-enhanced MARC)

سيمارك : مارك الموائم للمقررات الدراسية (سيمارك) .

Channel

القناة : موقع وب يرسل بشكل أوتوماتيكي معلومات محدثة للعرض الفوري أو عرضه حسب الطلب .

checksum

مجموع التدقيق : قيمة محسوبة اعتماداً على محتويات مجموعة بيانات وذلك لغرض الكشف عن فساد البيانات أو عدم صحتها .

CHIO (Cultural Heritage Information Online).

معلومات التراث الحضاري المباشرة : مبادرة ائتلاف التبادل الآلي للمعلومات المتحفية.

CIDOC (The International Committee for Documentation of the International council of Museums)

اللجنة الدولية للتوثيق التابعة للمجلس الدولي للمتاحف .

CIMI (Consortium for the Computer Interchange of Museum Information)

ائتلاف التبادل الآلي للمعلومات المتحفية .

Clickthrough measurement

قياس تتبع الروابط .

Code list

قائمة الرموز .

Codebook

أدلة الترميز أو التكويد .

Composite

تركيبة .

Compound element

عنصر مركب .

Content ratings

[تقديرات] المحتوى .

Content rules

قواعد المحتوى .

Continuing Resources

المصادر المستمرة .

Control field

حقل ضبط .

Controlled Vocabulary

لغة مقيدة .

Copy Cataloging

الفهرسة المنقولة .

CORC (cooperative online Resource Catalog)

كورك [الفهرس التعاوني للمصادر المباشرة].

Crosswalks

نقاط المقابلة أو التحويل البيني .

CSDGM (Content Standard for Digital Geospatial Metadata) ,

معيار محتوى ما وراء البيانات المكانية الرقمية .

CUSTARD (U.S./Canadian Standards Reconciliation Project).

مشروع توافق المعايير الأمريكية والكندية .

Darwin Core

داروين كور أو فئات داروين الأساسية .

ZBIG (Biological Implementors Group)

مجموعة المنفذين البيولوجيين .

Data field

حقل بيانات [حقل متغير].

Dataset.toe

ملف يعرف مجموعة من المحتوى الإلكتروني في صيغة إفيكت .

DCMI (Dublin Core Metadata Initiative)

مبادرة دبلن لما وراء البيانات .

DDI (Data Documentation Initiative)

مبادرة توثيق البيانات .

Deep web

الوب الخفية .

Denver Code

مجموعة فرعية من معيار محتوى ما وراء البيانات المكانية الرقمية للجنة الفيدرالية للبيانات الجغرافية .

Descriptive metadata

ما وراء البيانات الوصفية .

Digital signature

التوقيع الإلكتروني .

Directory

الدليل [في تسجيلة مارك].

DLF (Digital Libraries Federation)

اتحاد المكتبات الرقمية .

DOI (Digital Object Identifier)

معرف الكيان الرقمي .

Domain

مجال .

Domain-specific term

مصطلح مجال محدد .

DREL (Digital Rights Expression Language)

لغة التعبير عن الحقوق الرقمية .

DRM (Digital Rights Management)

إدارة الحقوق الرقمية .

DTD

معرف أو تعريف نوع الوثيقة .

Dumb down principle

مبدأ التهميش أو التجاهل .

EAD (Encoded Archival Description)

الوصف الأرشيفي المكود.

Ebind

أي بايند : مواصفة ما وراء بيانات بنائية .

EBX (Electronic Book eXchange)

أي بيكس أو مبادرة تبادل الكتاب الإلكتروني .

EDItEUR

إديتير : جماعة دولية اهتمت بتشجيع التجارة الإلكترونية في قطاعي الكتب والمسلسلات .

EFFECT (Exchange Format for electronic components and Texts)

إيفيكت ، صيغة تبادل المكونات والنصوص الإلكترونية .

Element refinement qualifier

مقيد [توضيح معنى] العنصر .

EML (Ecological Metadata language)

لغة ما وراء البيانات الأيكولوجية .

Empty element

العنصر الخالي .

Encoding scheme qualifier

مقيد خطة الترميز .

EPICS (EDItEUR Product Information Communication Standards)

إيبكس : معايير تبادل معلومات منتجات إديتير .

ESRI

شركة تطوير وتسوق برمجية نظم المعلومات الجغرافية.

Expression

أسلوب التعبير [في المتطلبات الوظيفية للتسجيلة الببليوجرافية] .

Extension scheme

خطة توسعة [في معيار ترميز ونقل ما وراء البيانات] .

METS (Metadata Encoding and Transmission Standard)

معيار ترميز ونقل ما وراء البيانات .

FGDC (Federal geographic data Committee)

اللجنة الفيدرالية للبيانات الجغرافية .

Finding aids

معينات البحث .

Fonds

الوحدة الأرشيفية الرئيسية : مصطلح أنجلو- كندي يستخدم للمجموعة الأرشيفية أو مجموعة السجلات .

FRBR

[المتطلبات الوظيفية للتسجيلات الببليوجرافية] . إف آر بي آر

GEM (the Gateway for the educational Materials)

جيم [بوابة المواد التعليمية].

GIF Graphics Interchange Format

جيف [صيغة التبادل البيني للرسومات أو الصور الرسومية].

GILS (Government Information Locator Service or Global Information Locator Service)

جيلس : خدمة محدد موقع المعلومات الحكومية أو خدمة محدد موقع المعلومات الكونية.

GIS (Geographic Information system(S))

نظم المعلومات الجغرافية .

Hidden web

الوب الخفية .

HTML (Hypertext Markup Language)

لغة ترميز النص الفائق .

HTTP (Hyper text transfer Protocol)

بروتوكول نقل الملفات الفائقة ، البروتوكول الذي هو أساس الوب .

ICONCLASS

نظام تصنيف موضوعي دولي للصور الفنية .

ICPSR (Inter-University Consortium for Political and Social Research)

ائتلاف الجامعات للأبحاث السياسية والاجتماعية .

IDF (the International DOI foundation)

المؤسسة الدولية تعريف الكيانات الرقمية .

IEC (International electrotechnical Commission)

اللجنة الدولية للتقنيات الإلكترونية .

IEEE (the Institute of Electrical and electronics Engineers)

جمعية المهندسين الكهربائيين ومهندسي الإلكترونيات تنطق i-triple-E, .

IETF (Internet engineering Task Force)

لجنة المهام الخاصة بهندسة الإنترنت .

ILS Integrated Library System

النظام المتكامل للمكتبة .

IMS (the IMS (Instructional Management Systems) Global Learning Consortium)

نظم الإدارة التعليمية لائتلاف التعليم الكوني.

<indecs> Interoperability of Data in E-Commerce Systems

إنديكس ، مشروع مولته المفوضية الأوربية خلال الفترة من ١٩٩٨ حتى ٢٠٠٠م .

Indicators

المؤشرات [في تسجيلة مارك].

Information package

حزمة معلومات [في نموذج OAIS] .

ISAD(G) (General International Standard archival Description)

التقنين الدولي للوصف الأرشيفي العام .

JPEG (Joint photography experts group)

جي بيج : طريقة أو أسلوب ضغط مقنن للصور الفوتوغرافية الرقمية .

KNB (Knowledge Network for Biocomplexity)

شبكة المعرفة للمركبات الحيوية .

LC Copy

نسخة مكتبة الكونجرس .

LCRI (Library of congress Rule Interpretations)

تفسيرات مكتبة الكونجرس لقواعد الفهرسة .

Leader

الفاتح [التمثيلات الأربع والعشرون الأولى من تسجيلة مارك].

Link analysis

تحليل الرابط أو الربط .

Linkage

التعبير عن العلاقات بين الكيانات .

Locator records

تسجيلات محددات المواضع .

LOM (Learning Object Metadata)

ما وراء بيانات الكيان التعليمي .

LTSC (the Learing Technology Standards Committee)

لجنة معايير تقنيات التعليم التابعة لجمعية مهندسي الإلكترونيات والمهندسين الكهربائيين IEEE .

Manifestation

وسائط العرض [في المتطلبات الوظيفية للتسجيلات الببليوجرافية] .

MARC

مارك (الفهرسة المقروءة آليًا) .

MARC٢١

مارك ٢١ .

MCF (Meta Content Framework)

إطار ما وراء المحتوى .

MD٥

مجموع التدقيق اللوغاريتمي .

MDA Museum Documentation Association

جمعية توثيق المتاحف أو [التوثيق المتحفي] .

Meta-metadata

وصف ما وراء البيانات .

Metadata scheme

خطة ما وراء بيانات .

Metalanguage

لغة متعدية أو لغة عن لغات، لغة تستخدم لوصف لغات أخرى .

METS (Metadata Encoding and Transmission Standard)

ميتس : معيار ترميز ما وراء البيانات ونقلها .

MIDI (Musical instrument Digital Interface)

واجهة رقمية للموسيقى الرقمية .

MIME type

أنواع مايم .

MIX (Metadata for Image in XML Schema),

ميكس : خطة بصيغة لغة الترميز الموسعة لتمثيل الدلالات الخاصة بمعيار المنظمة الوطنية لمعايير المعلومات والجمعية الدولية لإدارة المعلومات والصور لما وراء البيانات الفنية للصور الرقمية الثابتة .

MOA٢ (Making of America ١١) ,

تطور أمريكا الثانية، مشروع دعمه اتحاد المكتبات الرقمية لبناء مجموعة رقمية من المواد الأرشيفية .

MP٣

إم بي ثري : صيغة ضغط معيارية أو مقننة للصوت .

MPEG (Motion Picture Experts Group)

إم بيج : مجموعة خبراء الصور المتحركة ، منظمة تطور المعايير ذات الصلة بالفيديو والوسائط المتعددة .

Namespace

اسم مميز أو فريد .

NARA (the National Archives and Records Administration)

نارا : إدارة الأرشيفات والسجلات الوطنية .

NASA's Consultative Committee for Space Data Systems

اللجنة الاستشارية بوكالة ناسا لأنظمة البيانات الفضائية .

NBII (National Biological Information Infrastructure)

البنية التحتية الوطنية للمعلومات البيولوجية .

NEDLIB (Networked European Deposit Library)

مشروع مكتبة الإيداع الأوربية الشبكية .

NESSTAR

نيستار (أدوات ومصادر العلوم الاجتماعية الشبكية) .

Netcasting

البث عبر النت (ويسمى أيضاً البث الشبكي Webcasting).

NII (National Information Infrastructure)

البنية التحتية للمعلومات الوطنية .

NISO (National Information Standards Organization)

المنظمة الوطنية لمعايير المعلومات .

NLA

المكتبة الوطنية الأسترالية .

OAI (Open Archives Initiative)

مبادرة الأرشيفات المفتوحة .

OAIS (Open Archival Information System)

نموذج نظام المعلومات الأرشيفية المفتوح .

OCLC

مركز الحاسب الآلي للمكتبات على الخط المباشر .

ODRL (Open Digital Rights Language)

لغة الحقوق الرقمية المفتوحة .

OEB the (Open eBook publication structure)

بنية طبع أو نشر الكتاب الإلكتروني المفتوح .

OeBF (Open eBook forum)

منتدى الكتاب الإلكتروني المفتوح .

One – to- one principle

مبدأ "حالة بحالة" [في دبلن كور] .

ONIX (Guidelines for Online Information eXchange)

أونيكس .

Ontology

الأنطولوجيا .

Original order

الترتيب الأصلي .

PANADORA (preservation and accessing Networked Documentary resources of Australia)

بانادورا : حفظ المصادر الوثائقية وإتاحتها .

Preservation metadata

ما وراء بيانات الحفظ .

Profile

البروفايل .

Provenance

الأصل [تاريخ إنشاء وتملك المادة الأرشيفية].

PSD (Photoshop Document),

صيغة ملف الصور .

Public-key encryption

الترميز او التشفير بالمفتاح العام .

Qualifier

المقيد او المحدد [في دبلن كور].

RAD (Rules for Archival Description)

قواعد الوصف الأرشيفي .

RDF (Resource Description Framework)

إطار وصف المصدر .

Record group

مجموعة المحفوظات .

Reference name

الاسم المرجعي [في مواصفة أونيكس] .

REL (Rights Expression Language(S))

لغة التعبير عن الحقوق .

Remote access resources

المصادر التي يمكن الوصل إليها عن بعد أو المتاحة عن بعد .

Respect des fonds

مبدأ " احترام الوحدة الأرشيفية المتكاملة" [في التوثيق الأرشيفي] .

Rights metadata

ما وراء بيانات الحقوق .

RLG, (Research Libraries group)

مجموعة المكتبات البحثية .

RLIN, (Research Libraries Information Network).

شبكة معلومات المكتبات البحثية .

ROADS (Resource Organization And Discovery in Subject-based Services)

رودز : تنظيم المصادر واستكشافها في الخدمات المعتمدة على الموضوع .

RSS (RDF Site Summary or Rich Site Summary depending on the version)

ملخص الموقع لإطار عمل وصف المصدر أو ملخص الموقع الثري المعتمد على الإصدارة : أكثر المواصفات شهرة لتعريف القنوات على الوب .

SAA (Society of American Archivists)

جمعية الأرشيفيين الأمريكيين .

Schema

مخطط أو خطة .

SCORM (Sharable Content Object Reference Interface)

سكورم : واجهة مرجعية لكيان محتوى قابل للمشاركة .

SCSI (Small computer system interface)

واجهة نظام آلي صغيرة .

Semantic unit

وحدة دلالية [في المتطلبات الوظيفية للتسيجيلات الببليوجرافية FRBR].

Semantics

دلالات الألفاظ [تعريف لمعنى عناصر ما وراء البيانات ، في مقابل قواعد الترميز أو التكويد أو قواعد تمثيل قيم العناصر].

SGML (Standard Generalized Markup language)

لغة الترميز المعيارية العامة .

Short tag

تاج قصير [في مواصفة أونيكس].

SICI (Serial item and contribution identifier)

معرف مكونات الدورية وأجزائها .

SIP (Submission Information Package)

حزمة تقديم معلومات [في إطار نظام المعلومات الأرشيفي المفتوح].

SMIL (Synchronized Multimedia Integration Language)

لغة تكامل وسائط متزامنة .

Spamming

الإثقال أو التحميل المتعمد لصفحات الوب بالكلمات .

Spider

زواحف ، عناكب ، برنامج تستخدمه محركات بحث الإنترنت لبحث وجمع وتكشيف محتوى الوب ، [وتسمى أيضاً webcrawler].

SSA (Serial Storage Architecture)

البنية الفنية للتخزين المتسلسل .

Structural metadata

ما وراء البيانات التركيبية .

Subelement

عنصر فرعي .

Superwork

العمل الأكبر أو العمل فائق .

Surrogate

بديل .

Syntax

صيغة ، أو تركيبة .

Tag

تاج [في تسجيلة مارك] .

Tag library

مكتبة التيجان [دليل التيجان أو العناصر].

Technical metadata

ما وراء البيانات الفنية .

TEI (Text encoding Initiative)

مبادرة ترميز النص .

TGN (Thesaurus of Geographic Names)

مكنز جيتي للأسماء الجغرافية .

TIFF (Tag Image File Format)

تيف [صيغة ملف الصور] .

UKOLN (the united kingdom office for library and information networking)

مكتب المملكة المتحدة لمشابكة المكتبات والمعلومات .

ULAN (the Union List of Artist Names)

القائمة الموحدة لأسماء الفنانين .

UNICODE

يونيكود أحد معايير ISO/IEC للتمثيل الحرفي .

URI (Uniform resource Identifier)

المعرف الموحد للمصدر .

URL (Uniform resource Locator)

المحدد الموحد لموقع المصدر .

URN (Uniform resource Name)

الموحد للمصدر .

USMARC

مارك الأمريكي .

VADS (Visual Arts Data Service)

خدمة بيانات الفنون المرئية .

Voucher

فاوتشر [في مواصفة تبادل الكتاب الإلكتروني].

VRA (Core Visual Resources Association core Categories)

الفئات الأساسية لجمعية المصادر المرئية .

Work

العمل [في المتطلبات الوظيفية للتسجيلة الببليوجرافية] .

W٣C

اتحاد الوب .

Wrapper

الغلاف أو المغلف [في لغة الترميز الموسعة واللغة الموحدة العامة للترميز ،عنصر يشتمل بداخله على عنصر آخر ، أو مجموعة من العناصر (العناصر الفرعية)].

XML

لغة الترميز الموسعة .

XMP (eXtensible Metadata Platform)

برنامج أو بيئة عمل ما وراء البيانات القابلة للتوسع .

XrML (eXtensible rights Markup language)

اللغة الموسعة لترميز الحقوق .

Z tokens

رموز زد .

ZBIG (Biology Implementors Group)

مجموعة المنفذين البيولوجيين .

Z٣٩,٥٠ معيار زد ٣٩,٥٠.